

COMPUTER AND PARALLEL COMPUTER SYSTEM AND MESSAGE TRANSMISSION PROCESSING METHOD

Publication number: JP11134309 (A)

Publication date: 1999-05-21

Inventor(s): CHIBA NOZOMI; PATRICK HAMILTON

Applicant(s): HITACHI LTD

Classification:

- International: G06F15/173; G06F15/163; G06F15/16; (IPC1-7): G06F15/163

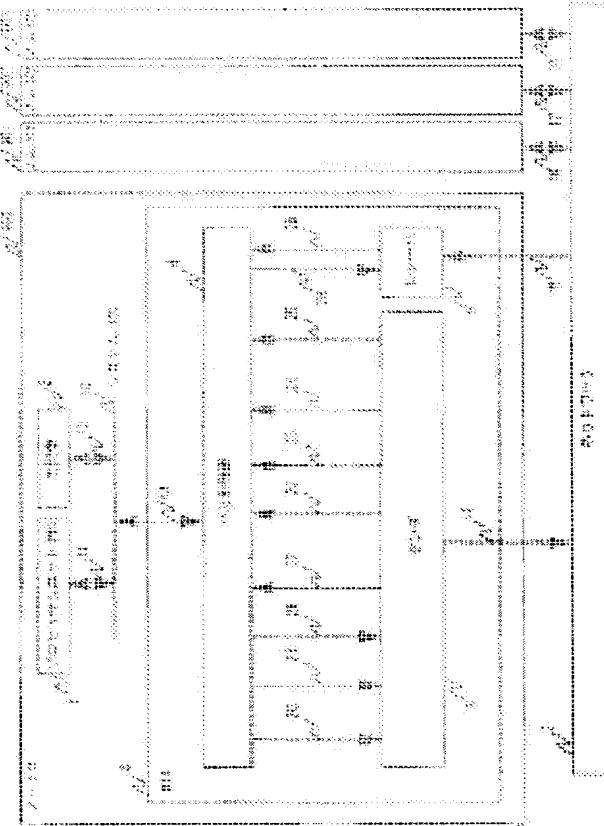
- European:

Application number: JP19970295242 19971028

Priority number(s): JP19970295242 19971028

Abstract of JP 11134309 (A)

PROBLEM TO BE SOLVED: To realize a parallel computer for a multi-user by interrupting and resuming a processing for transmitting the arbitrary number of messages in each node. **SOLUTION:** A sender part 5 of a network interface adapter(NIA) 3 functioning as an interface between a processor and a network in each node is provided with a means for detecting the establishment of a message transmission processing interruption condition, means for interrupting the message transmission processing for a message transmission unit when the message transmission processing interruption condition is detected, and means for storing control information (SMCW) for resuming the interrupted message transmission processing. The SMCW is processed for queuing in a main storage 2, and a processor 1 reads the SMCW from the main storage 2, and transmits it to the NIA 3. Thus, the interrupted message transmission processing can be resumed.



Data supplied from the esp@cenet database — Worldwide

(18) 日本国特許庁 (J P)

(12) 金門特許金額 (A)

(1)特許出願公開番号

特開平11-134309

(43) 公開日 平成11年(1999)5月21日

(51) Int.Cl. *

識別記号

F I

G O S E - 15/16

320 R

審査請求 未請求 請求項の数 8 ○ 1 (全 25 頁)

(21) 出願番号 特願平9-295242

(22) 出願日 平成9年(1997)10月28日

(71)出願人 000005108

株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地

(72) 素明者 千葉 重一
神奈川県秦野市堀山下1番地 株式会社日

(72) 発明者 パトリック・ハミルトン
神奈川県秦野市城山下1番地 株式会社日立製作所汎用コンピュータ事業部内

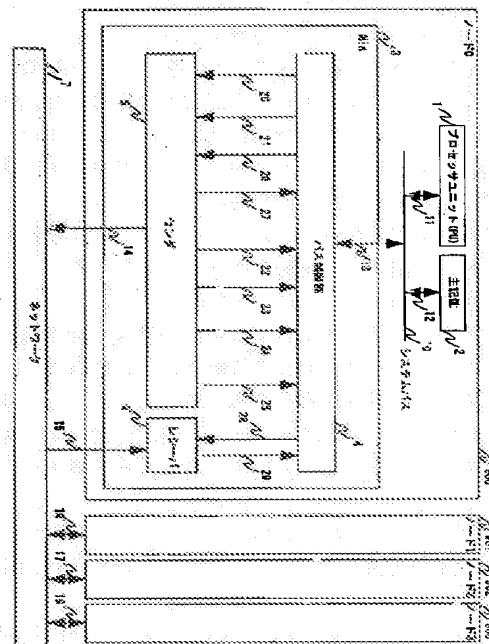
(74) 代理人：井理士 錦木 誠

(54) 【発明の名称】 計算機、並列計算機システム及びメッセージ送出処理方法

(57) **【愛約】**

【課題】 各ノード内で任意の数のメッセージ送出処理を中断および再開可能として、マルチユーザ向けの並列計算機を実現する。

【解決手段】 各ノード内で、プロセッサとネットワークとのインターフェースを司るネットワークインターフェースアダプタ (NIA) のセンダ部に、メッセージ送出処理中断条件の成立を検出する手段と、該メッセージ送出処理中断条件を検出した場合に該メッセージ送出処理をメッセージ送出単位に中断する手段と、該中断されたメッセージ送出処理を再開するための制御情報 (SMCW) を格納する手段を設ける。該SMCWは主記憶にキューイングし、プロセッサが該主記憶からSMCWを読み出してNIAに送付することにより、該中断されたメッセージ送出処理を再開する。



【特許請求の範囲】

【請求項1】 プロセッサと、該プロセッサで処理する命令列やデータを格納する主記憶と、ネットワークとのメッセージ送受信を司るネットワークインタフェースアダプタ（以下、NIAと称す）を具備する計算機において、

NIAが、メッセージ送出処理中断条件の成立を検出する手段と、前記メッセージ送出処理中断条件の成立を検出した場合に該メッセージ送出処理をメッセージ送出単位に中断する手段と、該中断されたメッセージ送出処理を再開する手段とを有することを特徴とする計算機。

【請求項2】 請求項1記載の計算機において、主記憶に、中断されたメッセージ送出処理を再開するための制御情報（以下、SMCWと称す）をキューイングし、プロセッサが該主記憶からSMCWを読み出してNIAに送付することにより、該NIAで中断されたメッセージ送出処理を再開することを特徴とする計算機。

【請求項3】 請求項1、2記載の計算機において、NIAが、メッセージ送出処理の経過時間を計測するタイマを有し、該経過時間があらかじめ定めた時間を超過したことをメッセージ送出処理中断条件とすることを特徴とする計算機。

【請求項4】 請求項1、2記載の計算機において、メッセージ送信のための制御情報（以下、TCWと称す）中にメッセージ送出中断ビットを設け、該ビットがメッセージ送出中断を示すことをメッセージ送出処理中断条件とすることを特徴とする計算機。

【請求項5】 請求項1乃至4記載の計算機において、NIAが、複数個のメッセージ送出処理要求を受け付ける手段を有し、メッセージ送出処理中断条件が成立し、かつ、メッセージ送出を行っているメッセージ送出処理要求の他にメッセージ送出処理要求がある場合に、メッセージ送出処理を中断することを特徴とする計算機。

【請求項6】 請求項1乃至5記載の計算機において、NIAが、SMCWを一時保持する手段と、メッセージ送出処理を中断したことをプロセッサに割り込みにより報告する手段を有し、該割り込みによりプロセッサが所定ソフトウェアを起動し、該ソフトウェアにより前記SMCWを読み出して主記憶にキューイングすることを特徴とする計算機。

【請求項7】 請求項1乃至6記載の構成の計算機が複数個、ネットワークを介して接続され、計算機間でメッセージ送受信を行う並列計算機システムであって、各計算機において任意の数のメッセージ送出処理の中止および再開を可能としたことを特徴とする並列計算機システム。

【請求項8】 複数の計算機がネットワークを介して接続され、計算機間でメッセージ送受信を行う並列計算機システムにおけるメッセージ送出処理方法であって、メッセージ送出処理を実行し、該メッセージ送出処理を

中断する必要があるか否か判定し、該メッセージ送出処理を中断する場合に該メッセージ送出処理をメッセージ送出単位に中断し、該中断したメッセージ送出処理を再開するための制御情報をキューイングすることを特徴とするメッセージ送出処理方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、複数のプロセッシングノードをネットワークで接続した並列計算機システムにおいて、特に一つのノード内でマルチユーザ環境を実現し、任意の数のメッセージ送出処理を中断および再開可能とする計算機及びその並列計算機システム、更にはメッセージ送出処理方法に関する。

【0002】

【従来の技術】一般的に並列計算機システムはプロセッシングノードをネットワークで接続した構成であり、プロセッシングノード間の通信を行うために、プロセッシングノード毎にメッセージの送受信を司るネットワークインタフェースアダプタ（NIA）が設けられている。メッセージ送受信のための制御情報はノード内の主記憶もしくはNIA内のレジスタに格納され、NIAはユーザまたはOSからのメッセージ送出要求を受信すると、該メッセージ送受信のための制御情報に基づいてメッセージ送受信を行う。メッセージ送受信のための制御情報が主記憶に格納される場合、特にメッセージ送出では、メッセージ送受信のための制御情報を複数個連結して、1回の送出要求で複数のメッセージ送出ができるようになる場合が多い。以降、特にメッセージ送信のための制御情報をTCWと呼び、該TCWを複数個連結したものをTCWチェインと呼ぶことにする。一般にメッセージ送出要求をNIAが受信した後でも、TCWをTCWチェインに連結することができる。

【0003】従来、並列計算機システムは主に科学技術計算分野に利用されてきた。科学技術計算では一人のユーザにより全てのノードで同一のジョブを実行する形式が多く、一つのノード内でマルチユーザ環境の実現はほとんどないため、従来の並列計算機システムでは、障害発生時は別として、通常のメッセージ送出処理途中で該メッセージ送出処理を中断および再開する機構は存在しなかった。

【0004】

【発明が解決しようとする課題】近年、並列計算機の用途がデータベースサーバやオンラインランザクションプロセスなどの非科学技術計算分野に広がるのに伴い、マルチユーザ環境の実現が必要となってきている。マルチユーザ環境下においては、ノード内のNIAが複数のユーザからのメッセージ送出要求を受け付ける。送出の順番はメッセージ送出要求をNIAが受け付けた順番であり、処理中のメッセージ送出処理が終了すると次のメッセージ送出要求の処理を開始する。ここで、非常に長

いTCWチェインが存在すると、そのメッセージ送出処理のために、他の多くのTCWやTCWチェインのメッセージ送出処理を長時間待たせ、システム全体の処理時間を延ばしてしまうことになる。また、ソフトウェアはTCWをTCWチェインにいくつでも連結できるため、他のメッセージ送出処理がいつまでも開始されない危険性がある。これらの問題を解決するためには、TCWとTCWとの間でメッセージ送出処理を中断し、かつ、中断したメッセージ送出処理を矛盾なく再開できる機構が必要である。しかし、NIAの論理規範などの制約で中断できるメッセージ送出処理数が限られていては、その限りを越えて中断できないメッセージ送出処理で前述の問題が発生することとなる。

【0005】本発明の目的は、メッセージ送出処理途中で任意にメッセージ送出処理を中断し、該中断されたメッセージ送出処理を矛盾なく再開することができ、かつ中断できるメッセージ送出処理に限りが無い、マルチユーザ向け計算機及びそれを使用した並列計算機システム、メッセージ送出処理方法を実現することにある。

【0006】
【課題を解決するための手段】本発明では、並列計算機システムを構成する各計算機のネットワークインターフェースアダプタ(NIA)に、メッセージ送出処理中断条件の成立を検出する手段と、該メッセージ送出処理中断条件の成立を検出した場合に該メッセージ送出処理をメッセージ送出単位に中断する手段と、該中断されたメッセージ送出処理を再開するための制御情報を格納する手段とを設ける。そして、ソフトウェアにより、該格納されたメッセージを再開するための制御情報を主記憶にキューイングするようにする。該中断されたメッセージ送出処理を再開するための制御情報をソフトウェアによりキューイングすることにより、NIA内に該中断されたメッセージ送出処理を再開するための制御情報をそのまま保持する必要がないため、該NIAの論理規範という物理的な制約に問らず任意の数のメッセージ送出処理を中断および再開することができる。

【0007】NIAは、メッセージ送出処理の経過時間を計測するタイマを有し、該経過時間があらかじめ定めた時間を超過したことをメッセージ送出処理中断条件とする。あるいは、TCW中にメッセージ送出中断ビットを設け、該ビットがメッセージ送出中断を示すことをメッセージ送出処理中断条件とする。

【0008】
【発明の実施の形態】以下、本発明の実施形態を図面に用い説明する。図1は本発明の並列計算機システムの一実施形態の全体構成図である。本システムは、計算などの処理を行う複数のノード0～3(900～903)をノード間でメッセージを転送するネットワーク7を介し信号線14～18で接続した構成である。ここで、信号線14はノード0(900)がネットワーク7へ送出す

るメッセージを、信号線15はノード0(900)がネットワーク7から受信するメッセージを伝える。ノード1～3(901～903)の各信号線16～18は、ノード0(900)の信号線14と15をひとつにまとめて表した信号線である。図1ではノード数を4個としているが、本発明はノード数が2個以上であれば有効である。各ノードの構成、動作は同じであるので、以下ではノード0(900)について詳述する。

【0009】ノード0(900)は、OSやユーザアプリケーションを実行するプロセッシングノード(以降PU)1、データを格納するための主記憶2、ネットワーク7とノード0(900)とのインターフェースを司るネットワークインターフェースアダプタ(以降NIA)3、PU1と主記憶2とNIA3とを接続するシステムバス10、システムバス10とPU1とを接続する信号線11、システムバス10と主記憶2とを接続する信号線12、システムバス10とNIA3とを接続する信号線13とから成る。

【0010】主記憶2は、図2に示すように、ポート番号で識別されそれぞれ送受信処理に必要なメッセージ制御用情報を格納した複数の仮想ポートと、ポート番号と仮想ポートの先頭アドレス(以降VPA)の対応テーブル(以降PAT)30を保持している。ここで、該仮想ポートの中に、メッセージ送信のための制御情報(TCW)を一つ以上連結したメッセージ送出処理情報チエイン(TCWチェイン)があり、仮想ポートの先頭からTCWチェインの先頭までのオフセットアドレス(以降先頭TCWアドレス)31が該仮想ポートの先頭(VPA)に置かれている。カーネルがPAT30を管理し、かつ仮想ポートの大きさをNIA3がレジスタに保持してあって、該NIA3がユーザにより発行された仮想ポート以外の領域をアクセスするリクエストを処理しないようにすることで、マルチユーザ間のプロテクションを保つことが出来る。TCW0(32)には、TCW0がTCWチェインの最終であるか否かを示す最終TCWビットと、メッセージ送出処理を中断するための送出中断ビットと、TCWチェインにおいて次のTCWのアドレスを示す次TCWアドレスと、ネットワーク7においてメッセージが通るルートを示すルーティング情報と、ノードが送出するメッセージの主記憶上のアドレスを示すメッセージ先頭アドレスと、ノードが送出するメッセージの長さを示すメッセージ長と、メッセージを受信するノードを示す宛て先ノードと、メッセージを受信したノードがメッセージ送信元を識別するためにメッセージに付加される送信元ノードとがある。後述するメッセージヘッダにはルーティング情報、メッセージ先頭アドレス、メッセージ長、宛て先ノード、送信元ノードが用いられる。TCWの内容はTCW1～3(33～35)についても同様である。ここで、TCW0(32)の次TCWアドレスはTCW1(33)のオフセットアドレス

(ポインタ)である。以下、TCW1(33)の次TCWアドレスはTCW2(34)のオフセットアドレスであり、TCW2(34)の次TCWアドレスはTCW3(35)のオフセットアドレスである。図2の場合、TCW0にチェインされたTCWの数は3個であるが、本発明はTCWの数によらず有効である。

【0011】図1に戻り、NIA3は、システムバス10とNIA3とのインターフェースを司るバス制御部4、ネットワーク7にメッセージを送出するセンダ5。ネットワーク7からのメッセージを受信するレシーバ6から成る。バス制御部4とセンダ5との間には、バス制御部4から送られるデータの種類を識別するためのデータ種類識別信号の信号線20と、該バス制御部4から送られるデータの信号線21と、ソフトウェアによりキューリングされる中断されたメッセージ送出処理を再開するための制御情報の信号線22と、該センダ5が送出要求を受け付けたか否かを示す送出要求受付結果の信号線23と、該センダ5がバス制御部4を介してメモリアクセスを行うために必要なデータアドレスとデータサイズとリードリクエストかライトリクエストかを区別する信号の信号線24と、メッセージ送出処理中断をPU1に報告する割り込み信号の信号線25と、中断されたメッセージ送出処理を再開するための制御情報の有効ビットをクリアするための信号線26と、メッセージ送出処理中断前の最後に処理したTCWの次TCWアドレスの信号線27がある。ここで、中断されたメッセージ送出処理を再開するための制御情報を以降SMCW(Suspended Message Control Ward)と呼び、その有効ビットをSMCW有効ビットと呼ぶことにする。これらの信号線20~27については、後でセンダ5の動作説明において詳しく述べる。また、バス制御部4とレシーバ6との間には、メッセージを受信するためバス制御部4からレシーバ6へ伝えられる信号の信号線28と、レシーバ6からバス制御部4へ伝えられる受信したメッセージの信号線29とがある。なお、バス制御部4およびレシーバ6の動作は、本発明と直接関係しないため詳細な説明は省略する。

【0012】次に、センダ5の構成例を図3で説明する。センダ5は、データ種類識別信号デコード40と、送出要求キュー制御部41と、DMAリクエスト制御部42と、TCWレジスタ制御部43と、送出バッファ制御部44と、送出要求キュー45と、TCWレジスタ46と、SMCWレジスタ47と、SMCWレジスタ有効ビット48と、データヘッダセレクタ49と、送出バッファ50と、内部信号線60~115からなる。

【0013】データ種類識別信号デコード40は、信号線20により伝えられるデータ種類識別信号から、信号線21によりバス制御部4から伝えられるデータがメッセージ送出要求か、VPAか、先頭TCWアドレスか、TCWか、メッセージデータか、タイムアウト時間かの

いずれかを識別し、それぞれの場合に応じて、送出要求信号、VPAリプライ信号、先頭TCWアドレスリプライ信号、TCWリプライ信号、データリプライ信号、タイムアウト時間セット信号を発行する。図3中、信号線60は信号線21のデータがメッセージ送出要求であることを示す送出要求信号を伝え、信号線61は信号線21のデータがVPAであることを示すVPAリプライ信号、先頭TCWアドレスであることを示す先頭TCWアドレスリプライ信号、信号線21のデータがネットワーク7へ送出するメッセージデータであることを示すデータリプライ信号などを伝え、信号線62はVPAリプライ信号、先頭TCWアドレスリプライ信号、TCWリプライ信号、さらに信号線21のデータがタイムアウト時間であることを示すタイムアウト時間セット信号などを伝え、信号線63はデータリプライ信号を伝え、信号線64もデータリプライ信号を伝える。

【0014】送出要求キュー45は、信号線21により伝えられるメッセージ送出要求を2つ以上保持するキューである。該送出要求キュー45は、後述の送出要求キュー制御部41の制御下でメッセージ送出要求をリード/ライトする。本実施の形態では、ソフトウェアがメッセージ送出を要求する場合、センダ5が信号線21を通して受信するデータは仮想ポートであるとする。図3中、信号線110は該送出要求キュー45から読み出されたメッセージ送出要求である仮想ポート番号を伝え、信号線112は該送出要求キュー45から読み出されたSMCWである仮想ポート番号を伝える。

【0015】送出要求キュー制御部41は、送出要求キュー45のライトアドレス、ライトネーブル、リードアドレスを管理し、当該センダ5がメッセージ送出処理中でない場合にメッセージ送出要求を受け付けた場合と、TCWの処理が終了したときに該送出要求キュー45にメッセージ送出要求がある場合に、送出開始要求信号を発行する。また、該送出要求キュー制御部41は、該送出要求キュー45にメッセージ送出要求が2つ以上あるか否かを検出し、2つ以上ある場合に送出待ち送出要求検出信号を発行する。図3中、信号線70は送出要求キュー45のライトアドレスとライトネーブルとリードアドレスを伝え、信号線71はメッセージ送出処理を開始する信号である送出開始要求信号を伝え、信号線72は該送出開始要求信号と、当該センダ5で送出処理を開始されていないメッセージ送出要求が送出要求キュー45にあることを示す送出待ち送出要求検出信号を伝える。なお、信号線23を介してバス制御部4へ、メッセージ送出要求を受け付けたか否かを示す送出要求受付結果を返すのも該送出要求キュー制御部41である。

【0016】DMAリクエスト制御部42はメッセージ送出処理のために必要なメモリアクセスリクエストを発行する。発行するリクエストはVPAリクエスト、先頭TCWアドレスリクエスト、先頭TCWリクエスト、デ

ータリクエスト、次TCWリクエスト、先頭TCWアドレスライトリクエストである。それぞれのリクエストについて、メモリアクセスリクエストのコマンド信号、リードリクエストかライトリクエストかを区別するリードライト信号、リクエストするデータのアドレスであるリクエストアドレス、リクエストするデータのサイズであるリクエストサイズを、信号線24を介してバス制御部4へ発行する。また、DMAリクエスト制御部42は、TCWレジスタ46のTCWに示されたデータ長分のデータリクエストが終了したことを示すリクエスト終了信号を発行する。図3中、信号線80はデータのリクエストが終了したことを示すリクエスト終了信号を伝え、信号線81はリクエストしたデータのサイズを示すリクエストサイズを伝える。

【0017】TCWレジスタ46は、DMAリクエスト制御部42のリクエストに対して信号線21により伝えられるVPA、先頭TCWアドレス及びTCWデータを保持するレジスタである。該TCWレジスタ46は、後述のTCWレジスタ制御部43の制御下でこれらデータをリード／ライトする。図3中、信号線111は該TCWレジスタ46から読み出されるVPA、先頭TCWアドレス、TCW中のメッセージ先頭アドレスとメッセージ長と最終TCWビットと送出中断ビットと次TCWアドレスなどを伝え、信号線113は該TCWレジスタ46から読み出されるTCW中のメッセージヘッダを伝える。

【0018】TCWレジスタ制御部43は、TCWレジスタ46のライトアドレス、ライトイネーブル、リードアドレスを管理し、他の制御部でTCWレジスタ46のデータが必要なときにリードアドレスを調整する。また、データ種類識別信号デコーダ40よりTCWリップライ信号が返ると、メッセージヘッダを生成し、送出バッファ制御部44を介して送出バッファ50へヘッダのライトを要求するヘッダライト信号を発行し、これが終了すると、ヘッダ完了信号をDMAリクエスト制御部42へ発行する。また、DMAリクエスト制御部42がリクエスト終了信号を発行すると、メッセージ送出処理が終了したのか、次のTCWの処理を開始するのか、メッセージ送出処理を中断するのかを判定する。そして、メッセージ送出処理が終了した場合は送出終了信号を、次のTCWの処理を開始する場合は次TCWリクエストを発行する。メッセージ送出処理を中断する場合は先頭TCWアドレスライトリクエストと、SMCWをSMCWレジスタ47へセットしつつSMCW有効ビット48をセットするSMCWセット信号と、信号線25を通してメッセージ送出処理を中断したことをPUへ報告する割り込み信号を発行する。さらに、該TCWレジスタ制御部43はメッセージ送出を開始してからの時間を計測するタイマ410を有し、該タイマ410の値がメッセージ送出処理を開始してからメッセージ送出処理中断条件が

成立するまでの時間（以降タイムアウト時間）を超過したこと（以降「タイムアウトした」という）を検出する。図3中、信号線90はメッセージヘッダを送出バッファ50にライトし終わったことを示すヘッダライト完了信号、TCWの処理が終了し次のTCWの取得を要求する次TCWリクエスト、TCWレジスタ46のTCW中の次TCWアドレスを主記憶上の仮想ポートの先頭に置かれた先頭TCWアドレスにライトするための先頭TCWアドレスライトリクエスト、データアドレスセット信号、データ長セット信号などを伝え、信号線91はTCWチェックの全てのTCWの処理を終了したこと若しくはメッセージ送出処理を中断したことを示す送出終了信号を伝え、信号線92はメッセージヘッダのライトを要求するヘッダライト信号を伝え、信号線93はTCWレジスタ46のライトアドレスとライトイネーブルとリードアドレスを伝え、信号線94はSMCWセットアドレスとライトイネーブルとリードアドレスを伝え、信号線94はSMCWセット信号を伝える。

【0019】送出バッファ50は、ネットワーク7へ送出するメッセージヘッダ、メッセージを一時格納するバッファである。該送出バッファ50は、後述の送出バッファ制御部44の制御下でメッセージヘッダ、メッセージをリード／ライトする。

【0020】送出バッファ制御部44は該送出バッファ50のライトアドレス、ライトイネーブル、リードアドレスを管理する。また、該送出バッファ50の空き容量を計算し、DMAリクエスト制御部42とTCWレジスタ制御部43へ伝える。図3中、信号線100は送出バッファ50の空き容量を伝え、信号線101は送出バッファ50のライトアドレスとライトイネーブルとリードアドレスを伝える。

【0021】SMCWレジスタ47はSMCWを保持するレジスタであり、SMCWレジスタ有効ビット48はSMCWレジスタ47の内容が有効であるか否かを示すビットである。図中、信号線114はSMCW有効ビット48の値をTCWレジスタ制御部43へ伝える。なお、SMCWレジスタ45のSMCWは、信号線25による割り込み報告を契機に、ソフトウェアにより信号線22を介して読み込まれる。

【0022】データヘッダセレクタ49は、TCWレジスタ46から信号線113を介して来るメッセージヘッダと信号線21を介してバス制御部4から来るメッセージデータとのどちらを送出バッファ50へライトするかをセレクトする。該データヘッダセレクタ49のセレクト信号はデータ種類識別信号デコーダ40から信号線63を介して伝えられるデータリップライ信号であり、メッセージデータが返る場合にメッセージデータを、それ以外ではメッセージヘッダをセレクトし、信号線115へ出力する。

【0023】まず、図1乃至図3と図8のメッセージ送

出処理フローチャートを基にして、本発明の一実施形態によるメッセージ送出処理と該メッセージ送出処理を中断および再開する処理を説明する。その後、図3の各制御部40～44の詳細について述べる。

【0024】図8に於いて、F100はメッセージ送出処理を開始する信号である送出開始要求信号を発行するか否かを判定するステージである。F101はVPAを、F102は先頭TCWアドレスを、F103はTCWを取得するステージである。F104は取得したTCWからメッセージヘッダを生成し送出バッファ50へライトするステージである。F105は該メッセージヘッダを送出バッファ50へライトし終えた後に、データリクエストとデータリクエストにより取得したメッセージデータを送出バッファ50へライトすることを並行して行うステージである。F106はデータリクエストが終了したTCWが、TCW cheインの最終TCWであるか否かを判定するステージである。F107はメッセージ送出処理を中断する条件が成立したか否かを判定するステージである。F108は最後に処理したTCWの次TCWアドレスを、主記憶2上の仮想ポートの先頭に置かれた先頭TCWアドレスへライトするステージである。F109はSMCWをSMCWレジスタ47へライトするステージである。F110はメッセージ送出処理中断をPU1に報告する割り込みを発生するステージである。以下、各ステージでの動作を説明する。

【0025】F100はメッセージ送出処理を開始する信号である送出開始要求信号を発行するか否かを判定するステージである。タイムアウトによりメッセージ送出処理を中断させたい場合は、メッセージ送出処理を開始する前にセンド5はタイムアウト時間の設定を行う。バス制御部4から、タイムアウト時間が信号線21を介して、また、該タイムアウト時間を設定することを示すデータ種類識別信号が信号線20を介してセンド5に伝えられると、データ種類識別信号デコード40が、送られて来たデータがタイムアウト時間であることを検出し、タイムアウト時間セット信号を信号線62を介してTCWレジスタ制御部43へ伝え、該TCWレジスタ制御部43は該タイムアウト時間セット信号により、信号線21上の該タイムアウト時間を取り込んで保持する。

【0026】メッセージ送出処理を開始する場合、PU1は仮想ポート番号をメッセージ送出要求としてNIA3に渡す。NIA3がメッセージ送出要求を受けると、バス制御部4はセンド5に、信号線20を介して送出要求受付要求を、信号線21を介してメッセージ送出要求の仮想ポート番号を伝える。データ種類識別信号デコード40は信号線20より送出要求受付要求を受信すると、送出要求キュー制御部41へ送出要求信号を信号線60を介して伝える。送出要求キュー制御部41は、送出要求キュー45に空きエントリがあるか否かを判定して、空エントリがあれば、信号線70を介して送出要求

キュー45へライトアドレスとライトイネーブルを伝えて、信号線21上の該メッセージ送出要求（仮想ポート番号）をキューイングし、空エントリがなければ、該メッセージ送出要求をキューイングしない。送出要求キュー制御部41は、メッセージ送出要求をキューイングしたか否かを送出要求受付結果信号として信号線23を介しバス制御部4経由でPU1へ伝える。次に、送出要求キュー制御部41は、送出要求信号を受け取ると、送出要求キュー45にメッセージ送出要求があり、メッセージ送出処理中でなければ、送出開始要求信号をDMAリクエスト制御部42とTCWレジスタ制御部43とへ信号線71、72を介して伝える。TCWレジスタ制御部43は、該送出開始要求信号により、メッセージ送出処理の経過時間を計測するタイマ（以降送出時間計測タイマ）410の計測を開始する。そして、図8のフローチャートはF100からF101ステージに進む。一方、該送出要求キュー45にメッセージ送出要求がない場合は、送出要求信号を受け取ってから、また、メッセージ送出処理中に送出要求信号を受け取った場合は、処理中のメッセージ送出処理が終了してから、送出要求キュー制御部41は送出開始要求信号を発行する。これらの場合、図8のフローチャートは、該送出開始要求信号を発行するまでF100ステージにとどまる。

【0027】また、F100ステージには、後述するF106ステージとF110ステージとから戻る場合がある。F106から戻った場合は、送出要求キュー45にメッセージ送出要求があればF101ステージへ進み、該送出要求キュー45に送出要求がなければF100にとどまる。F110から戻った場合は、メッセージ送出処理を中断する条件が成立した場合であり、該送出要求キュー45にメッセージ送出要求があることが該中断の条件であるため、必ずF101に進む。

【0028】F101ステージはVPAリクエストを発行しVPAを取得するステージである。DMAリクエスト制御部42は、送出要求キュー制御部41から信号線71を介して送出開始要求信号を受けると、送出要求キュー45から仮想ポート番号を信号線110を介して読み出し、VPA取得のためのVPAリクエストを信号線24を介してバス制御部4に発行する。本実施の形態では、バス制御部4がVPAリクエストとしてセンド5から仮想ポート番号を受け取ると、仮想ポート番号をPATのアドレスへ変換するものとする。該PATを参照して、仮想ポート番号に対応するVPAが主記憶2から読み出されると、バス制御部4は、信号線20を介してデータ種別識別信号としてVPA識別信号を、信号線21を介して該VPAを、センド5に伝える。データ種類識別信号デコード40は、信号線20を介してVPA識別信号を受信すると、信号線61、62にVPAリプライ信号を発行する。該VPAリプライ信号を信号線62を介して受信したTCWレジスタ制御部43は、信号線9

3を介してTCWレジスタ46へライトアドレス、ライトイネーブルを発行して、信号線21のVPAを該TCWレジスタ46へ保持する。さらに、TCWレジスタ制御部43は、信号線93を介してTCWレジスタ46へリードアドレスを発行し、該TCWレジスタ46から該VPAを読み出す。図8のフローチャートはF101からF102に進む。

【0029】F102ステージは、該VPAを用いて先頭TCWアドレスリクエストを発行し、先頭TCWアドレスを取得するステージである。本実施の形態に於いては、VPAと先頭TCWアドレスが格納されている主記憶上のアドレスは一致している。DMAリクエスト制御部42は、信号線61を介してVPAリプライ信号を受信すると、上記TCWレジスタ46から読み出されるVPAを信号線111を介して取り込み、該VPAをリクエストアドレスとして、先頭TCWアドレスリクエストを信号線24を介してバス制御部4に発行する。該VPA先に基づき、先頭TCWアドレスが主記憶2から読み出されると、バス制御部4は先頭TCW識別信号と該先頭TCWアドレスを信号線20、21を介してセンダ5に伝える。データ種類識別信号デコーダ40は、信号線20を介して先頭TCWアドレス識別信号を受信すると、信号線61、62に先頭TCWアドレスリプライ信号を発行する。ここで、該先頭TCWアドレスリプライ信号を受信したTCWレジスタ制御部43は、再び信号線93を介してTCWレジスタ46へライトアドレス、ライトイネーブルを発行して、信号線21の先頭TCWアドレスを該TCWレジスタ46に保持する。さらに、TCWレジスタ制御部43は、信号線93を介してTCWレジスタ46へリードアドレスを発行し、該TCWレジスタ46から該先頭TCWアドレスを読み出す。図8のフローチャートはF102からF103へ進む。

【0030】F103ステージは、該先頭TCWアドレスを用いて先頭TCWリクエストを発行し、TCWチキンの先頭のTCWを取得するステージである。DMAリクエスト制御部42は、先頭TCWアドレスリプライ信号を信号線61を介して受信すると、上記TCWレジスタ46から読み出される先頭TCWアドレスを信号線111を介して取り込み、該先頭TCWアドレスをリクエストアドレスとして、先頭TCWリクエストを信号線24を介してバス制御部4に発行する。該先頭TCWアドレスに基づきTCWデータが主記憶2から読み出されると、バス制御部4はTCW識別信号とTCWデータを信号線20、21を介してセンダ5に伝える。データ種類識別信号デコーダ40は、信号線20を介してTCW識別信号を受信すると、TCWリプライ信号をTCWレジスタ制御部43へ信号線62を介して伝える。TCWレジスタ制御部43は該TCWリプライ信号を受信すると、再び信号線93を介してTCWレジスタ46へライトアドレス、ライトイネーブルを発行して、信号線2

1のTCWデータを該TCWレジスタ46に保持する。

図8のフローチャートはF103からF104に進む。【0031】F104ステージは、該取得したTCWからメッセージヘッダを生成し送出バッファ50へライトするヘッダライトステージである。TCWレジスタ制御部43は、信号線93を介してTCWレジスタ46へ順次リードアドレスを発行して、該TCWレジスタ46からTCWの読み出し制御を開始し、該TCW中のメッセージヘッダ部分を読み出す際に、信号線92を介してヘッダライト信号を送出バッファ制御部44へ発行する。データヘッダセレクタ49は、セレクト信号であるデータリプライ信号が信号線63を介してアサートされていないため、信号線113を介して入力されるTCWレジスタ46の出力を選択し、信号線114に出力する。送出バッファ制御部44は、信号線92を介してヘッダライト信号を受信すると、信号線101を介して送出バッファ50へライトアドレスとライトイネーブルを発行し、該送出バッファ50へ該TCWレジスタ46の出力のメッセージヘッダ部分をライトする。該送出バッファ50へライトされたメッセージヘッダはネットワーク7へ送出される。ここで、送出バッファ制御部44は送出バッファ50の空き容量を計算して、TCWレジスタ制御部43へ信号線100を介して伝えている。TCWレジスタ制御部43は、送出バッファ50の空き容量がメッセージヘッダのライト単位よりも大きい場合に上記ヘッダライト信号を発行する。また、TCWレジスタ制御部43は、TCWレジスタ46からTCW中のデータ先頭アドレス(メッセージ先頭アドレス)とデータ長(メッセージ長)を読み出す際に、データアドレスセット信号とデータ長セット信号信号線90を介してDMAリクエスト制御部42へ発行する。該DMAリクエスト制御部42は、該データアドレスセット信号とデータ長セット信号とを受信すると、TCWレジスタ46から読み出されるメッセージデータ先頭アドレスとメッセージデータ長を信号線111を介して取り込んで保持する。メッセージヘッダのライトが終了すると、TCWレジスタ制御部43はDMAリクエスト制御部42へヘッダ完了信号を信号線90を介して発行する。図8のフローチャートはF104からF105に進む。【0032】F105ステージは、メッセージヘッダをネットワーク7へ送出し終えた後に、データリクエストと該データリクエストにより取得したメッセージデータを送出バッファ50へライトすることを並行して行うステージである。ヘッダ完了信号を契機として、DMAリクエスト制御部42はデータアドレス、リクエストサイズ、リードリクエストを信号線24を介してバス制御部4へ順次発行し、メッセージデータの取得するためのメモリアクセスを開始する。信号線20、21より、メッセージデータ識別信号、メッセージデータが到着すると、データ種類識別信号デコーダ40はデータリプライ

信号を信号線61、63、64を介して、DMAリクエスト制御部42とデータヘッダセレクタ49と送出バッファ制御部44とに伝える。DMAリクエスト制御部42は、リクエストサイズとデータアドレスとを加算し、また、リクエストサイズを残りデータ長から減算して保持しており、データリプライ信号を受信すると、それらの値を用いて次のメモリアクセスリクエストを発行する。また、DMAリクエスト制御部42は、リクエストサイズ(データサイズ)を信号線81を介して送出バッファ制御部44に伝える。データヘッダセレクタ49は、データリプライ信号を受信すると、信号線21のメッセージデータを選択し、信号線115へ出力する。送出バッファ制御部44は、データリプライ信号により、信号線101の送出バッファ50のライトアドレスをデータサイズ分更新し、ライトイネーブルをアサートして、信号線115のメッセージデータを送出バッファ50へライトする。DMAリクエスト制御部42はTCWに示されたメッセージ長分のデータのメモリアクセスリクエストを終了すると、リクエスト終了信号を信号線80を介してTCWレジスタ制御部43へ伝える。図8のフローチャートはF105からF106に進む。

【0033】F106はデータリクエストが終了したTCWが、TCWチェインの最終TCWであるか否かを判定するステージである。TCWレジスタ制御部43は、信号線80を介してリクエスト終了信号を受信すると、信号線93を介してTCWレジスタ46にリードアドレスを発行し、信号線111に読み出される該TCW中の最終TCWビットから、該データリクエストが終了したTCWがTCWチェインの最終TCWであるか否かを判定する。最終TCWである場合、メッセージ送出処理が終了したとみなして、送出時間計測タイマ410の計測を停止およびリセットし、また、送出要求キュー制御部41へ送出終了信号を信号線91を介して伝える。この時、図8のフローチャートはF100ステージに戻る。一方、最終TCWでない場合は、F107ステージへ進む。

【0034】F107はメッセージ送出処理を中断する条件が成立したか否かを判定するステージである。TCWレジスタ制御部43は、信号線72を介して受け取る送出待ち送出要求検出信号と、信号線114を介して受け取るSMCW有効ビット48と、タイマ410でタイマアウトが発生したか若しくは信号線111に読み出される該TCW中のメッセージ中断ビットが中断を示す値であるかを判定し、TCWチェインの次のTCWの処理を開始するのか、メッセージ送出処理を中断するのかを判定する。TCWチェインの次のTCWの処理を開始する場合には、TCWレジスタ制御部43はDMAリクエスト制御部42へ信号線90を介して次TCWリクエストを伝える。この時、図8のフローチャートはF103へ戻り、DMAリクエスト制御部42は、信号線111

に読み出される該TCW中の次TCWアドレスをリクエストアドレスとして、再びTCWリクエストを信号線24を介してバス制御部4に発行する。一方、メッセージ送出処理を中断する場合には、TCWレジスタ制御部43は、送出時間計測タイマ410の計測を停止およびリセットし、信号線91を介して送出終了信号を送出要求キュー制御部41へ発行し、図8のフローチャートはF108へ進む。

【0035】F108は最後に処理したTCWの次TCWアドレスを、主記憶2上の該当仮想ポートの先頭に置かれた先頭TCWアドレスへライトするステージである。TCWレジスタ制御部43は、信号線93を介してリードアドレスをTCWレジスタ46を発行して、該TCWレジスタ46が保持しているVPAと該TCW中の次TCWアドレスの読み出しを制御し、DMAリクエスト制御部42へ先頭TCWアドレスライトリクエストを発行する。DMAリクエスト制御部42は、TCWレジスタ46から読み出されたVPAを信号線111を介して取り込み、該VPAを主記憶上のライトアドレスに用いて、信号線24を介してバス制御部4へメモリライトアクセスリクエストを発行する。また、TCWレジスタ46から読み出された次TCWアドレスをライトデータとして、信号線27を介してバス制御部4へ伝える。図8のフローチャートはF108からF109へ進む。

【0036】F109は送出要求キュー45に残っているSMCWをSMCWレジスタ47へライトするステージである。送出要求キュー制御部41は、信号線91を介して送出終了信号を受信すると、送出要求キュー45に信号線70を介してリードアドレスを発行し、該送出要求キュー45に残っているSMCW(仮想ポート番号)を信号線112に出力する。TCWレジスタ制御部43は、信号線94を介してSMCWセット信号を発行して、送出要求キュー45から信号線112へ出力される仮想ポート番号をSMCWレジスタ47へセットし、また、該SMCWセット信号によりSMCW有効ビット48を有効にセットする。図8のフローチャートはF109からF110へ進む。

【0037】F110はメッセージ中断をPU1に報告する割り込みを発生するステージである。TCWレジスタ制御部43は、メッセージ送出処理中断を示す割り込み信号を信号線25を介してバス制御部4に発行する。図8のフローチャートはF100へ戻る。

【0038】PU1は該メッセージ送出処理中断割り込み信号を検出すると、SMCWレジスタ47からSMCW(仮想ポート番号)を信号線22を介して読み出し、信号線26を介してSMCWクリア信号を伝えることによりSMCW有効ビット48をクリアする。PU1が読み出した該SMCWとしての仮想ポート番号は、ソフトウェアにより主記憶2上の特定領域に順番にキューイングされ、該中断されたメッセージ送出処理を再開する場

合は、該キューイングした順番に仮想ポート番号を読み出し、再びメッセージ送出要求としてNIA3に伝え。これにより、図8のフローチャートは再びF100から開始する。なお、中断されたメッセージ送出処理を再開する契機は、本発明では特定しないが、例えば、主記憶2上の該特定領域を所定時間毎に読みに行くなどが考えられる。

【0039】なお、図8の処理シーケンスでは、メッセージ送出処理が中断した時、送出要求キュー45からSMCWを一旦SMCWレジスタ47へライトし、その後、割込みを契機に該SMCWを主記憶2の特定領域にキューイングするましたが、メッセージ送出処理の中断時、送出要求キュー45からSMCWをあらかじめ主記憶2内に設けた特定領域にライトし、その後に該主記憶2の別の領域にソフトウェアによりキューイングすることも可能である。

【0040】次に、図4乃至図7により送出要求キュー制御部41、DMAリクエスト制御部42、TCWレジスタ制御部43、送出バッファ制御部44の詳細について説明する。

【0041】図4は送出要求キュー制御部41の詳細図である。送出要求キュー制御部41は、メッセージ送出要求をセンダ5が受け付けたか否かを示す値を保持する受付結果レジスタ200と、送出要求キュー45のライトアドレスとリードアドレスとを保持するライトアドレスレジスタ201とリードアドレスレジスタ202と、インクリメント回路210～211と、比較器220～221と、ライトアドレスの値とリードアドレスの値とから送出要求キュー45の送出要求数を計算する送出要求数計算回路230と、信号の立ち上がりを検出する立ち上がり検出回路231と、AND回路240～242と、OR回路250と、信号線260～271とから成る。

【0042】送出要求キュー制御部41が信号線60を介して送出要求信号を受信すると、比較回路220の出力が“0”であれば、該メッセージ送出要求を受け付ける。この場合、AND回路241は信号線261に“1”を出力し、送出要求キュー45のライトイネーブルを信号線70を介してアサートし、インクリメント回路210の出力を信号線262を介してライトアドレスレジスタ201にセットし、該ライトアドレスレジスタ201の値をライトアドレスとする。一方、送出要求信号を受信したときに比較回路220の出力が“1”であれば、該メッセージ送出要求を受け付けずに、AND回路240の出力“1”を信号線260を介して受付結果レジスタ200にセットする。ここで比較器220は、信号線268を介して伝えられる送出要求数計算回路230が計算した送出要求キュー45内のメッセージ送出要求数と信号線269を介して伝えられる送出要求キュー45の面数とを入力とする。本実施の形態に於いては

送出要求キュー45を4面と仮定しており、図4では信号線269は4の値で示している。該メッセージ送出要求数と送出要求キュー45の面数が一致した場合、つまり送出要求キュー45がFULLの場合に比較器220の出力は“1”となる。本実施の形態に於いては、受付結果レジスタ200の値は、メッセージ送出要求が受け付けられた場合を“0”、メッセージ送出要求が受け付けられなかった場合を“1”とする。PU1からのメッセージ送出要求に対し、該受付結果レジスタ200を送出要求受付結果として信号線23、バス制御部4を介してPU1へ伝える。

【0043】送出待ち送出要求検出信号は比較器221により発行される。比較器221は、信号線268を介して送出要求数計算回路230の送出要求数と、信号線271を介して1の値とを入力とする。比較器221では該送出要求数と1とを比較し、該送出要求数が2以上の場合は送出待ち送出要求検出信号として“1”を信号線72に出力する。即ち、信号線72の出力“1”は、送出要求キュー45にまだ送出処理の開始されていないメッセージ送出要求があることを示している。

【0044】送出開始要求信号は立ち上がり検出回路231により発行される。ここで、OR回路250は信号線268を介して入力された送出要求数の各ビットのORをとり、該送出要求数が1以上の場合に“1”を出力し、送出要求キュー45にメッセージ送出要求が有ることを示す。AND回路242は、信号線266を介して入力されるOR回路250の出力と信号線91を介して入力される送出終了信号の否定とをANDし、信号線267を介して立ち上がり検出回路231に伝える。立ち上がり検出回路231は、信号線267の信号立ち上がりを検出し、送出要求数が0から1に変化する場合、つまり空の送出要求キュー45にメッセージ送出要求がキューイングされた場合と、送出要求数が1以上であり信号線91の送出終了信号が0→1→0と変化する場合、つまりメッセージ送出処理が終了したときに送出待ちのメッセージ送出要求がある場合とに、送出開始要求信号として“1”を信号線71、72に出力する。

【0045】メッセージ送出処理が終了または中断され、信号線91を介して送出終了信号が発行されると、リードアドレスレジスタ202は信号線264を介して入力されるインクリメント回路211の出力を保持し、該リードアドレスレジスタ202の値をリードアドレスとして信号線70に発行する。

【0046】図5はDMAリクエスト制御部42の詳細図である。DMAリクエスト制御部42は、リクエストしたデータのリプライ待ちであることを示すデータリップライ待ちビット300と、データリクエストを行っていることを示すデータリクエストビット301と、リクエストするデータのアドレスを保持するデータアドレスレジスタ302と、リクエストすべきデータの残量を保持

する残りデータ長レジスタ303と、OR回路310～313と、AND回路320～321と、比較器330と、加算器340と、減算器341と、データ先頭アドレスと加算器340の出力とをセレクトするデータアドレスセレクタ350と、データ長と減算器341の出力をセレクトするデータ長セレクタ351と、リクエストサイズとアドレスを出力するリクエストサイズ・アドレス選択回路352と、信号線360～388とからなる。

【0047】DMAリクエスト制御部42の全体の動作を説明する前に、まずリクエストサイズ・アドレス選択回路352の動作を説明する。該リクエストサイズ・アドレス選択回路352は、信号線61を構成する信号線363、364により伝えられる先頭TCWアドレスリプライ信号とVPAリプライ信号、信号線71により伝えられる送出開始要求信号、信号線90を構成する信号線369、370により伝えられる先頭TCWアドレスライトリクエストと次TCWリクエストなどをセレクトし、信号として入力し、また、信号線110により伝えられる仮想ポート番号、信号線111により伝えられるVPAと先頭TCWアドレスと次TCWアドレス、信号線377により伝えられるデータアドレスレジスタ302の値、信号線381により伝えられる残りデータ長レジスタ303の値などを被セレクト信号として入力し、信号線371にリクエストアドレスを、信号線372にリクエストサイズを出力するセレクタ回路である。ここで、信号線371、372は信号線24の一部を構成し、さらに信号線372は信号線81にもなっている。該リクエストサイズ・アドレス選択回路352は、信号線71から送出開始要求信号を受信すると、リクエストアドレスとして信号線110の仮想ポート番号を選択し信号線371へ出力する。VPAのサイズはシステムで固定長であるため、該リクエストサイズ・アドレス選択回路352で生成する。信号線61を構成するの信号線364からVPAリプライ信号を受信すると、リクエストアドレスとして信号線111のVPAを選択し、先頭TCWアドレスのリクエストサイズを生成し、それぞれ信号線371、372へ出力する。信号線61を構成する信号線363より先頭TCWアドレスリプライ信号を受信すると、リクエストアドレスとして信号線111の先頭TCWアドレスを選択し、TCWのリクエストサイズを生成し、それぞれ信号線371、372へ出力する。先頭TCWアドレスのリクエストサイズとTCWのリクエストサイズとは、VPAのサイズと同様固定長であるため、該リクエストサイズ・アドレス選択回路352で生成する。信号線363の先頭TCWアドレスリプライ信号と、信号線364のVPAリプライ信号と、信号線71の送出開始要求信号と、信号線369の先頭TCWアドレスライトリクエストと、信号線370の次TCWリクエストとのどれもが受信されない場合、該リクエスト

サイズ・アドレス選択回路352は、データアドレスレジスタ302の値をリクエストアドレスとして選択し、信号線371へ出力する。また、この場合のリクエストサイズとしては、残りデータ長レジスタ303の値と1回のリードリクエストで取得できる最大のデータ長とを比較し、小さい方の値を信号線372へ出力する。該1回のリードリクエストで取得できる最大のデータ長は固定長であり、リクエストサイズ・アドレス選択回路352で生成する。

【0048】次に、DMAリクエスト制御部42全体の動作を説明する。DMAリクエスト制御部42が信号線24、バス制御部4を介して主記憶2へ発行するリクエストはVPAリクエスト、先頭TCWアドレスリクエスト、先頭TCWリクエスト、データリクエスト、次TCWリクエスト、先頭TCWアドレスライトリクエストである。主記憶アクセスのための信号線24は信号線384、369、371、372で構成される。

【0049】信号線71を介して送出開始要求信号を受信すると、DMAリクエスト制御部42はVPAリクエストを発行する。該送出開始要求信号は信号線71を介してOR回路312に入力され、OR回路312は信号線384にメモリアクセスリクエストとして“1”を出力する。また、リクエストサイズ・アドレス選択回路352は信号線110の仮想ポート番号を選択して信号線371にリクエストアドレスとして出力し、さらにVPAのサイズを生成して信号線372にリクエストサイズとして出力する。信号線369はリードリクエストとライトリクエストを区別する信号線であるが、該信号線369は信号線90の一つで、TCWレジスタ制御部43が発行する先頭TCWアドレスライトリクエストを伝えられるもので、いまの場合は“0”である。本実施の形態では、信号線369が“0”的場合をリードリクエスト、“1”的場合をライトリクエストを示すものとする。なお、ライトリクエストは後述する先頭TCWアドレスライトリクエストの場合のみである。

【0050】VPAが主記憶2から読み出され、信号線61を構成する信号線364を介してVPAリプライ信号を受信すると、該DMAリクエスト制御部42は先頭TCWアドレスリクエストを発行する。リクエストサイズ・アドレス選択回路352は信号線111で伝えられるVPAをリクエストアドレスとして選択し、先頭TCWアドレスサイズをリクエストサイズとして生成し、信号線371、372に出力する。信号線384は“1”、信号線369は“0”である。

【0051】信号線61を構成する信号線363を介して先頭TCWアドレスリプライ信号を受信すると、該DMAリクエスト制御部42は先頭TCWリクエストを発行する。この場合、信号線371のリクエストアドレスは信号線111で伝えられる先頭TCWアドレスをリクエストサイズ・アドレス選択回路352が選択したもの

であり、信号線372のリクエストサイズはリクエストサイズ・アドレス選択回路352で生成した値である。信号線384は“1”、信号線369は“0”である。【0052】主記憶2から読み出されたTCWが返ると、TCWレジスタ制御部43はTCWレジスタ46に該TCWを保持する。TCWレジスタ制御部43は該TCWレジスタ46からのTCWの読み出し制御で、データアドレスセット信号とデータ長セット信号を、それぞれ信号線90を構成する信号線385、386を介して該DMAリクエスト制御部42へ伝える。信号線385のデータアドレスセット信号によりデータアドレスセレクタ350は信号線111により、TCWレジスタ46から伝えられる該TCW中のデータアドレス（メッセージ先頭アドレス）を選択し、また、OR回路310の出力は“1”であるため、データアドレスレジスタ302は該データアドレスを信号線376を介して保持する。また、信号線386のデータ長セット信号により、データ長セレクタ351は信号線111によりTCWレジスタ46から伝えられる該TCW中のデータ長（メッセージ長）を選択し、また、OR回路313の出力は“1”であるため、残りデータ長レジスタ303は該データ長を信号線380を介して保持する。

【0053】信号線90を構成する信号線362を介してヘッダ完了信号を受信すると、データリクエストビット301が“1”にセットされる。このとき残りデータ長レジスタ303の値は“0”ではないため、該残りデータ長レジスタ303の各ビットのORをとるOR回路311の出力は“1”である。また、信号線100を介して伝えられる送出バッファ50の空き容量が信号線372のリクエストサイズ以上であるときに、比較器330は“1”を出力する。信号線360を介してデータリプライ待ちビット300の出力“0”的否定と、信号線365を介してデータリクエストビット301の出力“1”と、信号線366を介してOR回路311の出力“1”と、信号線367を介して比較器330の出力“1”とがAND回路321に入力されると、AND回路321は信号線368を介して“1”をOR回路312に出力し、信号線384を介してメモリアクセスリクエストをバス制御部4を発行する。このときリクエストサイズ・アドレス選択回路352はデータアドレスレジスタ302の値をリクエストアドレスとして選択し、また、リクエストサイズを生成し、信号線371、372に出力する。信号線369は“0”である。

【0054】AND回路321は信号線368を介してデータリプライ待ちビット300を1にセットするため、上記メモリアクセスリクエストは1マシンサイクルのみアサートされる。このときTCWリプライ信号が来ないため、データアドレスセレクタ350は信号線378を介してデータアドレスレジスタ302の値とリクエストサイズ・アドレス選択回路352が出力するリクエ

ストサイズとを加算器340で加算した値を選択しており、また、信号線368を介してOR回路310の出力は“1”であるため、データアドレスレジスタ302は次のデータリクエストのデータアドレスとして信号線376を介して伝えられるデータアドレスセレクタ350の出力を保持する。データ長セレクタ351は、信号線383を介して伝えられる残りデータ長レジスタ303の値からリクエストサイズ・アドレス選択回路352が输出するリクエストサイズを減算器341で減算した値を選択しており、また、OR回路313の出力は同様に“1”であるため、残りデータ長レジスタ303は次の残りデータ長として信号線380を介して伝えられるデータ長セレクタ351の出力を保持する。

【0055】発行したデータリクエストに対するデータ（メッセージ）が返ると、信号線61を構成する信号線361を介してデータリプライ信号がアサートされ、データリプライ待ちビットが“0”にリセットされる。これにより、残りデータ長が0でなく、かつ、送出バッファ空き容量がリクエストサイズ以上であれば、AND回路321の出力が“1”となり、該DMAリクエスト制御部42は再びデータリクエストを発行する。DMAリクエスト制御部42は該データリクエストを残りデータ長が“0”になるまで行う。残りデータ長レジスタ303の値が0になると、OR回路311の出力は“0”となり、データリクエストは発行されない。最後のデータが到着すると、データリプライ待ちビット300が“0”にリセットされるため、AND回路320の出力が“0”から“1”へ変化する。立ち上がり検出回路342は該AND回路320の出力を信号線379を介して受取り、立ち上がりを検出すると、リクエスト終了信号を信号線80を介して発行する。また、該リクエスト終了信号によりデータリクエストビット301は0にリセットされる。

【0056】信号線90を構成する信号線370を介して次TCWリクエストを受信すると、該DMAリクエスト制御部42は次TCWメモリアクセスリクエストを発行する。該次TCWリクエストを受信すると、OR回路312の出力が“1”となり、信号線384を介してメモリアクセスリクエストが発行され、リクエストサイズ・アドレス選択回路352は信号線111を介してTCWレジスタ46から読み出した次TCWアドレスを選択し、またTCWのサイズを生成し、信号線371、372に出力する。

【0057】信号線90を構成する信号線369を介して先頭TCWアドレスライトリクエストを受信すると、該DMAリクエスト制御部42は先頭TCWアドレスライトのメモリアクセスリクエストを発行する。該TCWアドレスライトリクエストを受信すると、OR回路312の出力が“1”となり、信号線384を介してメモリアドレスアクセスリクエストが発行され、また、信号線

369の“1”によりライトリクエストが示される。リクエストサイズ・アドレス選択回路352は、信号線111を介して伝えられるVPAをライトアドレスとして選択し、信号線371に出力する。

【0058】図6はTCWレジスタ制御部43の詳細図である。TCWレジスタ制御部43は、タイムアウト時間を持するタイムアウト時間レジスタ400と、タイムアウトしたことを保持するタイムアウトビット401と、メッセージ送出処理を開始してからの経過時間を計測するタイマ(以降送出時間計測タイマ)410と、TCWをTCWレジスタ46に保持する際に使用するTCWカウンタ411と、メッセージヘッダを送出バッファ50に何回ライトしたかを示すヘッダライトカウンタ412と、TCWレジスタ46のリードアドレスとライトイネーブルとライトアドレスとを生成し、かつTCWレジスタ46に保持したTCWからデータアドレスとデータ長とをDMAリクエスト制御部42へセットするためのデータアドレスセット信号及びデータ長セット信号を生成するリードアドレスライトイネーブルライトアドレス生成回路(以降REW生成回路)413と、比較器420~423と、OR回路430~433と、AND回路440~444と、信号線362~473とから成る。

【0059】TCWレジスタ制御部43は、TCWレジスタ46のリード/ライト制御と、メッセージヘッダのライトと、タイムアウトの検出を行い、またTCWの処理が終了した場合にメッセージ送出処理が終了したのか、次のTCWを処理するのか、メッセージ送出処理を中断するのかを判定する。

【0060】まず始めにTCWレジスタ46のリード/ライト制御について説明する。さらにTCWレジスタ46のリード/ライト制御に関連して、図9を用いてTCWレジスタ46を、図13の表を用いてREW生成回路413を、図10を用いてメッセージを説明する。

【0061】図9は、TCWレジスタ46に信号線21を介してVPA、先頭TCWアドレス、TCWが保持される様子を示している。本実施の形態では、TCWレジスタ46は8エントリで構成され、VPAはエントリ0に、先頭TCWアドレスはエントリ1に、TCWはエントリ2~7に保持される。図2に示したように、エントリ2のA、Bはそれぞれ最終TCWビットと送出中断ビットを表わしている。

【0062】図13はREW生成回路413の入出力対応表である。REW生成回路413は、信号線62を構成する信号線459のVPAリプライ信号と信号線460の先頭TCWアドレスリプライ信号と信号線461のTCWリプライ信号、信号線411を介してのTCWカウンタ411の値、信号線92のヘッダライト信号、信号線466を介してのヘッダライトカウンタ412の値、信号線80のリクエスト終了信号、信号線90の一

部を構成する信号線471の次TCWリクエスト信号と信号線369の先頭TCWアドレスライトリクエストなどを入力とし、信号線93のTCWレジスタ46のライトアドレスとライトイネーブルとリードアドレス、信号線90の一部を構成する信号線385の該データアドレスセット信号と信号線386のデータ長セット信号などを出力とする。図13で、ライトアドレスとリードアドレスの列の数字はTCWレジスタ46のエントリを示し、ライトイネーブルが“1”的場合にライトアドレスが示すエントリにデータをライトする。また、データ先頭アドレスセット信号が“1”である場合にDMAリクエスト制御部42のデータアドレスレジスタ302にデータ先頭アドレス(メッセージ先頭アドレス)を、データ長セット信号が“1”である場合にDMAリクエスト制御部42の残りデータ長レジスタ303にデータ長(メッセージ長)を、それぞれセットする。ここで、「*」は“don't care”でありどんな値でも構わない。さらに入力のVPAリプライ信号と、先頭TCWアドレスリプライ信号と、TCWリプライ信号と、ヘッダライト信号と、次TCWリクエスト信号と、先頭TCWアドレスライトリクエストとはそれぞれ背反であり、同時に2つ以上の信号がアサートされることはない。また、本実施の形態では、TCWリプライデータは主記憶から数ビートで返ることを仮定しており、TCWリプライの各ビート毎に該TCWリプライ信号がアサートされるとする。以下、図13の表の順に説明する。

【0063】信号線62を構成する信号線459を介してVPAリプライ信号を受信すると、REW生成回路413により、TCWレジスタ46のエントリ0にVPAをライトするためライトアドレスが0、ライトイネーブルが“1”となる。同時に該VPAを用いて先頭TCWアドレスリクエストを発行するため、リードアドレスもエントリ0となる。信号線460を介して先頭TCWアドレスリプライ信号を受信した場合も同様に、ライトアドレスが1、ライトイネーブルが“1”、さらに該先頭TCWアドレスを用いて先頭TCWリクエストを発行するため、リードアドレスも1となる。信号線461を介してTCWリプライ信号を受信すると、TCWカウンタ411の値が更新されて、TCWレジスタ46のライトアドレスが更新し、図9のエントリ順にTCWをライトしていく。センダ5は該TCWのライト中は該TCWレジスタ46のデータを使わないため、リードアドレスは“*”である。TCWをTCWレジスタ46にライトし終わると、メッセージヘッダを送出バッファ50にライトし始める。図10は送出バッファ50にライトされたメッセージの内容を示している。メッセージはメッセージヘッダとメッセージデータとにより構成され、メッセージヘッダには、メッセージについての情報であるルーティング情報と、メッセージ先頭アドレスと、メッセージ長と、メッセージの宛て先ノードと、メッセージの

送信元ノードとが格納されている。これらは全てTCWに含まれている。また、メッセージデータには、DMAリクエスト制御部42がデータリクエストにより取得したデータが格納される。図13に戻って、ヘッダライト信号では、TCWレジスタ46のリードアドレスは図10に示したメッセージヘッダを送出バッファ50にライトするように更新される。リクエスト終了信号を受信する場合は、TCWレジスタ46から最終TCWビットAと送出中断ビットBを読み出すため、リードアドレスを2とする。更に、次TCWリクエスト信号を発行する場合は、リクエストアドレスがTCWの次TCWアドレスであるため、やはりリードアドレスを2とする。また、先頭TCWアドレスライトリクエストの場合は、ライトリクエストアドレスがVPA、ライトデータがTCWの次TCWアドレスであるため、リードアドレスは0と2になる。

【0064】図6に戻ってTCWカウンタ411を説明する。該TCWカウンタ411は、信号線62を構成する信号線461を介してTCWリプライ信号を受信するごとに更新するカウンタである。図13より、本実施の形態ではTCWをTCWレジスタ46にライトし終わるときにはTCWカウンタ411の値は5になっている。比較器421は信号線462を介して入力される該TCWカウンタ411の値と信号線463を介して入力される値5とを比較し、両者が一致した場合“1”を信号線464に出力し、該TCWカウンタ411をリセットする。

【0065】次に、メッセージヘッダの送出バッファ50へのライトについて説明する。前述のように、TCWをTCWレジスタ46にライトし終わると、比較器421は信号線464に1を出力する。これによりOR回路432の出力が“1”となり、ヘッダライトカウンタ412が1に更新され、該ヘッダライトカウンタ412の出力の各ビットのORをとるOR回路433は“1”を信号線468に出力する。また、比較器422は信号線100により伝えられる送出バッファ50の空き容量と、信号線472のメッセージヘッダを送出バッファ50へライトする単位であるヘッダライト単位とを比較し、該送出バッファ空き容量が該ヘッダライト単位よりも大きい、つまりヘッダをライトしても送出バッファ50がオーバーフローしない場合に、比較器422は信号線469に“1”を出力する。本実施の形態ではヘッダライト単位を4としてある。OR回路433と比較器422との出力が両方も“1”である場合に、信号線468、469を介してAND回路444は“1”を出力し、ヘッダライト信号を信号線92に出力する。また、該ヘッダライト信号によりOR回路432は“1”を出力し、ヘッダライトカウンタ412を更新する。比較器423は該ヘッダライトカウンタ412の値と信号線467を介して伝えられる5とが、ヘッダのライトが完了

して一致した場合に、信号線362にヘッダ完了信号として“1”を出力し、ヘッダライトカウンタ421をリセットする。

【0066】次にタイムアウトを検出する部分を説明する。信号線62を構成する信号線456を介してタイムアウト時間セット信号を受信すると、タイムアウト時間レジスタ400は信号線21を介して伝えられるタイムアウト時間を保持する。送出時間計測タイマ410は、信号線72を構成する信号線457により送出開始要求信号を受信すると、送出時間の計測を開始し、後述の信号線91により伝えられる送出終了信号を受信することで送出時間の計測を止めタイマ410をリセットする。比較器420は、タイムアウト時間レジスタ400と送出時間計測タイマ410の値をそれぞれ信号線450、451を介して受取り、両者が一致すると“1”を出力し、信号線452を介してタイムアウトビット401を“1”にセットする。該タイムアウトビット401は信号線91の送出終了信号により“0”にクリアされる。

【0067】最後に、信号線80を介してリクエスト終了信号を受信した場合に、メッセージ送出処理が終了したのか、次のTCWを処理するのか、メッセージ送出処理を中断するのかを判定する部分を説明する。信号線80のリクエスト終了信号が“1”であり、かつTCWレジスタ46から信号線111を構成する信号線455を介して読み出された最終TCWビットが“1”、つまりTCWチェインの最終TCWのリクエストを終了した場合には、AND回路440は出力“1”を信号線470を介してOR回路431に伝え、該OR回路431は信号線91を介して送出終了信号を発行する。また、信号線80のリクエスト終了信号が“1”であり、かつ信号線455の最終TCWビットが“0”であり、かつ信号線72構成する信号線473を介して伝えられる送出待ち送出要求検出信号が“1”であり、かつ信号線114を介して伝えられるSMCW有効ビット48が“0”でSMCWレジスタ47にSMCWをライトできる場合であり、かつ信号線458を介して伝えられる、タイムアウトビット401の値と信号線111を構成する信号線454を介して伝えられる送出中断ビットとを入力とするOR回路433の出力が“1”的場合に、AND回路442はメッセージ送出処理中断を示す割り込み信号を信号線25を介して、また、SMCWレジスタセット信号を信号線94を介して、先頭TCWアドレスライトリクエストを信号線369を介して、それぞれ発行する。さらに、OR回路431は、このときも出力“1”を送出終了信号として信号線91を介して発行する。信号線80のリクエスト終了信号が“1”であり、かつ信号線455の最終TCWビットが“0”であり、かつ信号線114のSMCW有効ビット信号の否定と信号線473の送出待ち送出要求検出信号とOR回路430の出力“1”との3つの値のANDがAND回路441で成立

しない場合には、AND回路443が信号線370を介して次TCWリクエストを発行する。

【0068】図7は送出バッファ制御部44の詳細図である。送出バッファ制御部44は送出バッファ50のライトアドレスを示すライトアドレスレジスタ500と、送出バッファ50のリードアドレスを示すリードアドレスレジスタ501と、OR回路510～511と、加算器520と、インクリメント回路521と、ライトアドレスレジスタ500とリードアドレスレジスタ501との値から送出バッファ50の空き容量を計算する空き容量計算回路522と、送出バッファ50にライトするヘッダまたはデータのサイズをセレクトするサイズセレクタ523と、信号線530～536とから成る。

【0069】OR回路510は信号線92を介してヘッダライト信号を、または信号線64を介してデータリップライ信号を受信すると、信号線533を介して送出バッファ50のライトイネーブルを出力する。セレクタ523は信号線92を介して入力されるヘッダライト信号が“1”的場合に信号線536を介して入力されるヘッダのライト単位を、信号線92のヘッダライト信号が“0”的場合に信号線81を介して入力されるリクエストサイズをセレクトする。本実施の形態ではヘッダのライト単位を4と仮定しており、図7では4で表している。加算器520は、ライトアドレスレジスタ500の値と信号線537を介して伝えられるサイズセレクタ523の出力を入力とし、信号線533のライトイネーブルが“1”的場合、該ライトアドレスレジスタ500は信号線530を介して伝えられる加算器520の出力を新しいライトアドレスとして保持する。信号線534は該ライトアドレスレジスタ500の値をライトアドレスとして出力する。リードアドレスレジスタ501は、信号線532を介して伝えられるOR回路511の出力が“1”的場合に、信号線531を介して伝えられるインクリメント回路521の出力を保持する。信号線535は該リードアドレスレジスタ501の値をリードアドレスとして出力する。ここでOR回路511は、空き容量計算回路522の出力を入力とし、該空き容量計算回路522の値が0でない場合、つまり送出バッファ50にネットワークへ送出すべきデータがある場合に“1”を出力する。空き容量計算回路522は、ライトアドレスレジスタ500の値とリードアドレスレジスタ501の値とを入力とし、信号線100を介して送出バッファ50の空き容量を出力する。

【0070】以上、本発明の実施の形態におけるメッセージ送出処理とメッセージ送出処理を中断および再開する動作と、センドラの詳細構成を説明した。

【0071】本発明の実施の形態に示したシステムについて、TCWレジスタ制御部43内のタイムアウト時間レジスタ400にタイムアウト時間を設定しておけば、多くのTCWが連結された長いTCWチェインを有する

送出要求（送出要求1と呼ぶ）のメッセージ送出処理を行っており、かつ次の少ないTCWが連結された短いTCWチェインを有する送出要求（送出要求2）が送出処理の開始を待っている場合でも、タイムアウト時間が経過するとタイムアウトが発生し、そのときに処理を行っているTCWの処理が終了した時点で送出要求1のメッセージ送出処理を中断し送出要求2の処理を開始する。そのため送出要求1が動的にTCWチェインを連結し続けるような場合でも、次の送出要求2がいつまでも待たされることとなり、システム全体の処理時間が延びるのを防ぐことと、ユーザが不当に長い時間メッセージ送出処理を行い他のユーザに迷惑をかけるのを防ぐことが出来る。さらにメッセージ送出処理を中断されたメッセージ送出要求のSMCWは、SMCWレジスタ47から読み出され、ソフトウェアにより主記憶上の所定領域にキューイングされ、該NIA3の内部に留まらないため、中断できる送出要求に限りが無い。そのため、例えば送出要求1～10が長いTCWチェインを有し、それ以降の送出要求11～20が短いTCWチェインを有しているような場合でも、送出要求1～10はタイムアウトが発生しメッセージ送出処理が中断されるため、送出要求11～20のメッセージ送出処理を不当に長い時間待たせることなく開始することが出来る。またTCWには該送出処理中断ビットを設けてある。TCWチェインが長くなるため、システム全体の処理時間を延ばしたり、他のユーザに迷惑をかけることがあらかじめわかっている場合には、TCWチェイン中の適当なTCWの該送出処理中断ビットを“1”に設定しておくことで、メッセージ送出処理を中断することができ、他のユーザに迷惑をかけずに済む。

【0072】本発明の実施の形態に於いては、メッセージ送出処理を中断する条件として、TCWレジスタ制御部43内の送出時間計測タイマ410の値がタイムアウト時間を超過した場合とTCWの送出中断ビットが“1”である場合との両方を採用しているが、どちらか一方のみを採用することでもよい。また本発明の実施の形態に於いては、NIAがメッセージ送出処理の中斷をPUに割り込みにより報告したが、他の方法としては、(1)NIA内にPUがポーリングするレジスタを設け、該レジスタにメッセージ送出処理の中斷を示す値を格納する方法と、(2)主記憶上にPUがポーリングする領域を設け、該領域にNIAがメッセージ送出処理の中斷を示す値を格納する方法などが考えられる。さらに本発明の実施の形態に於いては一つのTCWにつき1メッセージであったが、メッセージをより細かいパケットに分割することでもよい。

【0073】次に、本発明の他の実施の形態について説明する。今まででは、メッセージ送出を要求する場合、センドラがメッセージ送出要求として受信するデータは仮想ポート番号であったが、左程ユーザ間のプロトクショ

ンを保証しなくとも良い場合には、メッセージ送出要求として先頭TCWアドレスを渡す方法もある。この場合には、図2で示した仮想ポート、PATが不要であり、また、VPAリクエスト、先頭TCWアドレスリクエスト、先頭TCWアドレスライトリクエストが不要となる。それに伴いデータ種類識別信号デコード40、DMAリクエスト制御部42、TCWレジスタ制御部43の構成、動作が簡単になる。また、SMCWはTCWレジスタ46に保持された次TCWアドレスであり、SMCWを伝える信号線はTCWレジスタ46とSMCWレジスタ47との間で接続される。

【0074】図11に、この実施の形態のメッセージ送出処理フローチャートを示す。F200はメッセージ送出処理を開始する信号である送出開始要求信号を発行するか否かを判定するステージである。F201は先頭TCWリクエストを発行してTCWを取得するステージである。F202は取得したTCWからメッセージヘッダを生成し送出バッファ50へ送出するステージである。F203は該メッセージヘッダを送出バッファ50へライトし終えた後に、データリクエストと該データリクエストにより取得したメッセージデータを送出バッファ50へライトすることと並行して行うステージである。F204はデータリクエストが終了したTCWが、TCWチェインの最終TCWであるか否かを判定するステージであり、最終TCWであればF200に戻り、最終TCWでなければF205に行く。F205はメッセージ送出処理を中断する条件が成立したか否かを判定するステージであり、中断条件が成立しなければF200に戻り、中断条件が成立すればF206に行く。F206はTCWレジスタに保持されているSMCW(次TCWアドレス)をSMCWレジスタへライトするステージである。F207はメッセージ中断をPUに報告する割り込みを発生するステージであり、割り込み処理後、F200へ戻る。各ステージで行う処理は基本的に先の実施の形態の場合と同じである。

【0075】更に、TCWチェインを考慮しないシステムに於いては、メッセージ送出要求として直接TCWを渡す方法もある。この場合、送出要求キュー45にTCWが格納されるためTCWレジスタ46が不要になる。また、必要なリクエストはデータリクエストのみであり、それに伴いデータ種類識別信号デコード40、DMAリクエスト制御部42、TCWレジスタ制御部43の構成、動作が更に簡単になる。本実施の形態のシステムでは、TCWのメッセージをより細かなパケットに分割することで、メッセージ送出処理の中断ができる。また、SMCWはTCWであるが、送出要求として渡されたTCWと比べて、送出したパケットの分だけデータ先頭アドレスと残りデータ長が更新されたものとなる。

【0076】図12に、この実施の形態のメッセージ送出処理フローチャートを示す。F300はメッセージ送

出処理を開始する信号である送出開始要求信号を発行するか否かを判定するステージである。F301はメッセージ送出要求として取得したTCWからメッセージヘッダを生成し送出バッファ50へライトするステージである。F302はデータリクエストとデータリクエストにより取得したメッセージデータを送出バッファ50へライトすることと並行して行うステージである。F303はデータリクエストが終了か否かを判定するステージであり、終了であればF300に戻り、そうでなければF304に行く。F304はメッセージ送出処理を中断する条件が成立したか否かを判定するステージであり、中断条件が成立しなければF300に戻り、中断条件が成立すればF305に行く。F305はSMCW(TCW)をSMCWレジスタへライトするステージである。F306はメッセージ中断をPUに報告する割り込みを発生するステージであり、割り込み処理後、F300へ戻る。なお、本実施の形態でパケットのデータ領域の大きさは、1回のデータリクエストで取得できるデータ量よりも小さいと仮定した。

【0077】

【発明の効果】本発明によれば、任意の数のメッセージ送出処理を中断及び再開できる並列計算機システムを実現し、該並列計算機システムをマルチユーザ環境下で使用する場合、一人のユーザがメッセージ送出処理を占有することを防止できる。

【図面の簡単な説明】

【図1】本発明の並列計算機システムの実施の形態の全体構成図である。

【図2】主記憶上にあるポート番号・仮想ポート先頭アドレス対応テーブル(PAT)及び仮想ポートとメッセージ送出制御情報(TCW)チェインの例を示す図である。

【図3】ノード内のネットワークインタフェースアダプタのセンドアの構成図である。

【図4】センドア内の送出要求キュー制御部の詳細図である。

【図5】センドア内のDMAリクエスト制御部の詳細図である。

【図6】センドア内のTCWレジスタ制御部の詳細図である。

【図7】センドア内の送出バッファ制御部の詳細図である。

【図8】本発明による一実施の形態におけるメッセージ送出処理シーケンスのフローチャートである。

【図9】センドア内のTCWレジスタの詳細図である。

【図10】センドア内の送出バッファにライトされたメッセージの詳細図である。

【図11】本発明による他の実施の形態におけるメッセージ送出処理シーケンスのフローチャートである。

【図12】本発明による更に他の実施の形態におけるメ

メッセージ送出処理シーケンスのフローチャートである。

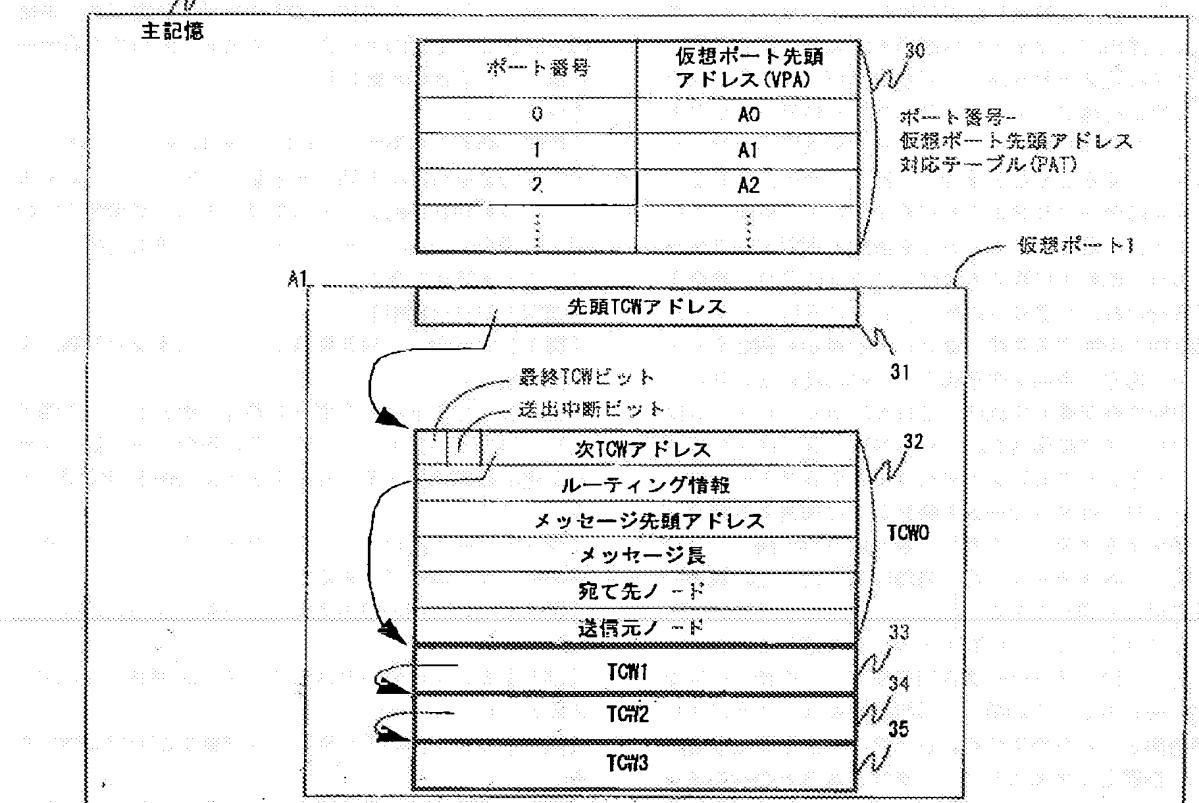
【図13】TCWレジスタ制御部内のリードアドレス・ライトイネーブル・ライトアドレス(REW)生成回路の入出力対応表である。

【符号の説明】

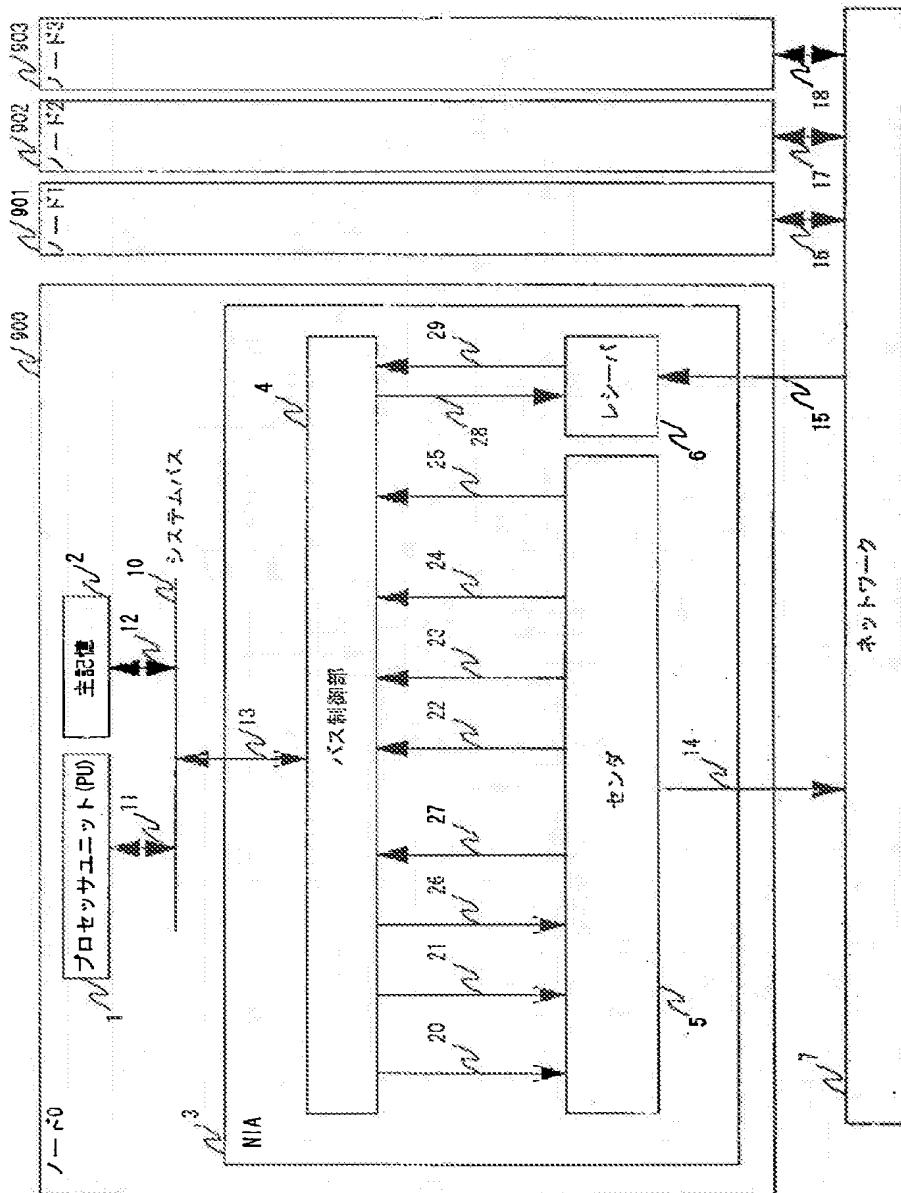
- 1 プロセッサユニット
 - 2 主記憶
 - 3 ネットワークインターフェースアダプタ (NIA)
 - 4 バス制御部
 - 5 センダ
 - 6 レシーバ
 - 7 ネットワーク

爲了這一個問題，我們在這裏就不得不說說「N」了。

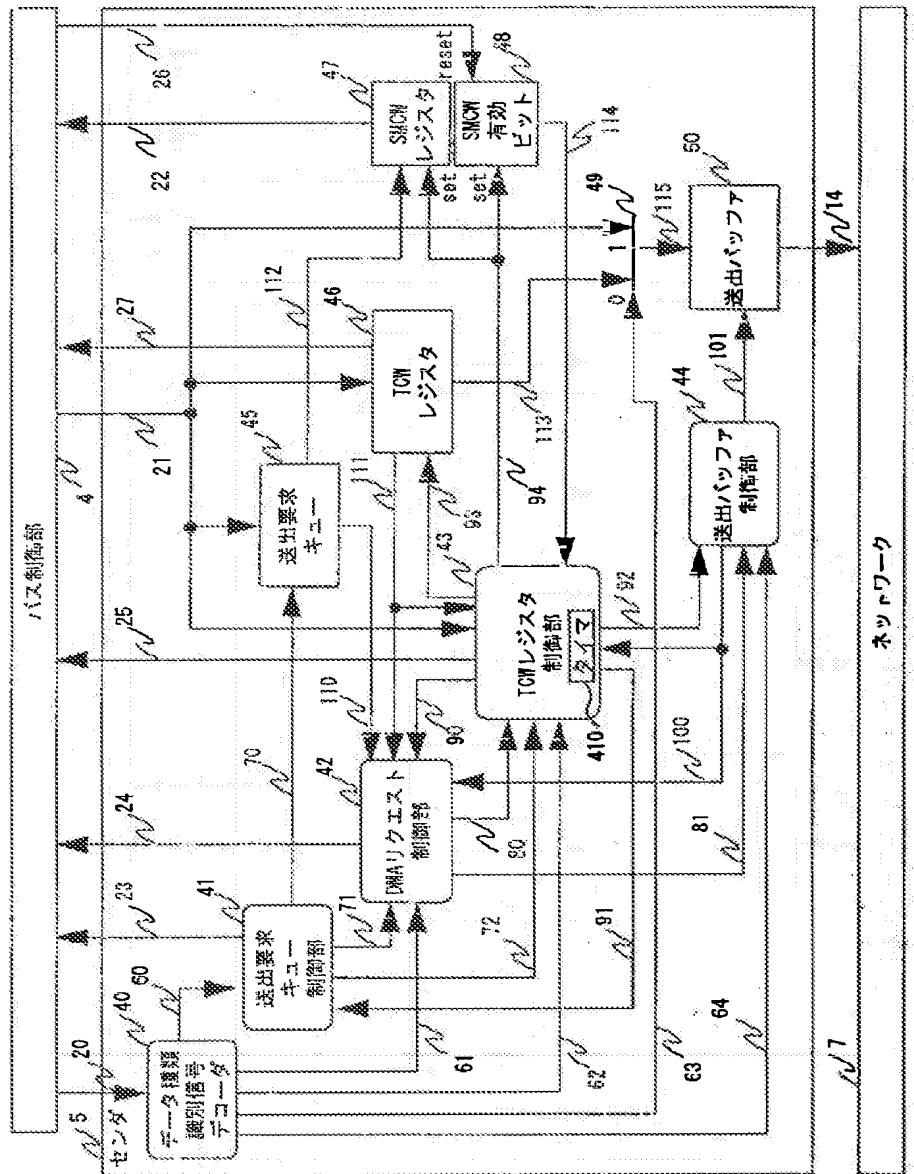
- 4 0 データ種類識別信号デコーダ
- 4 1 送出要求キュー制御部
- 4 2 DWAリクエスト制御部
- 4 3 TCWレジスタ制御部
- 4 4 送出バッファ制御部
- 4 5 送出要求キュー
- 4 6 TCWレジスタ
- 4 7 SMCWレジスタ
- 4 8 SMCW有効ビット
- 4 9 セレクタ
- 5 0 送出バッファ



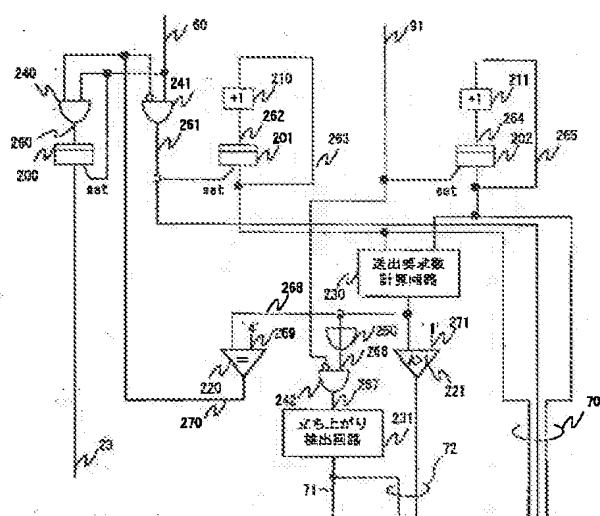
【図1】



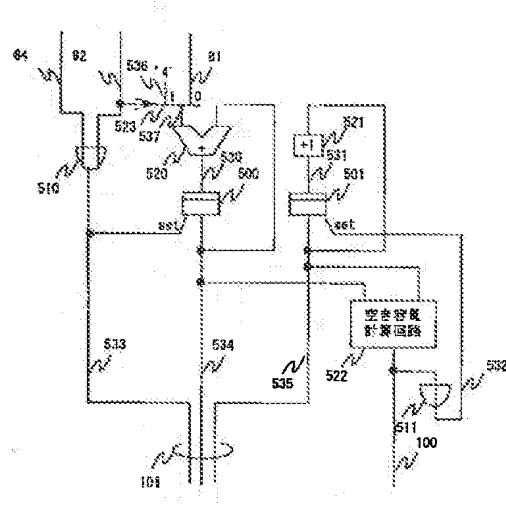
[图3]



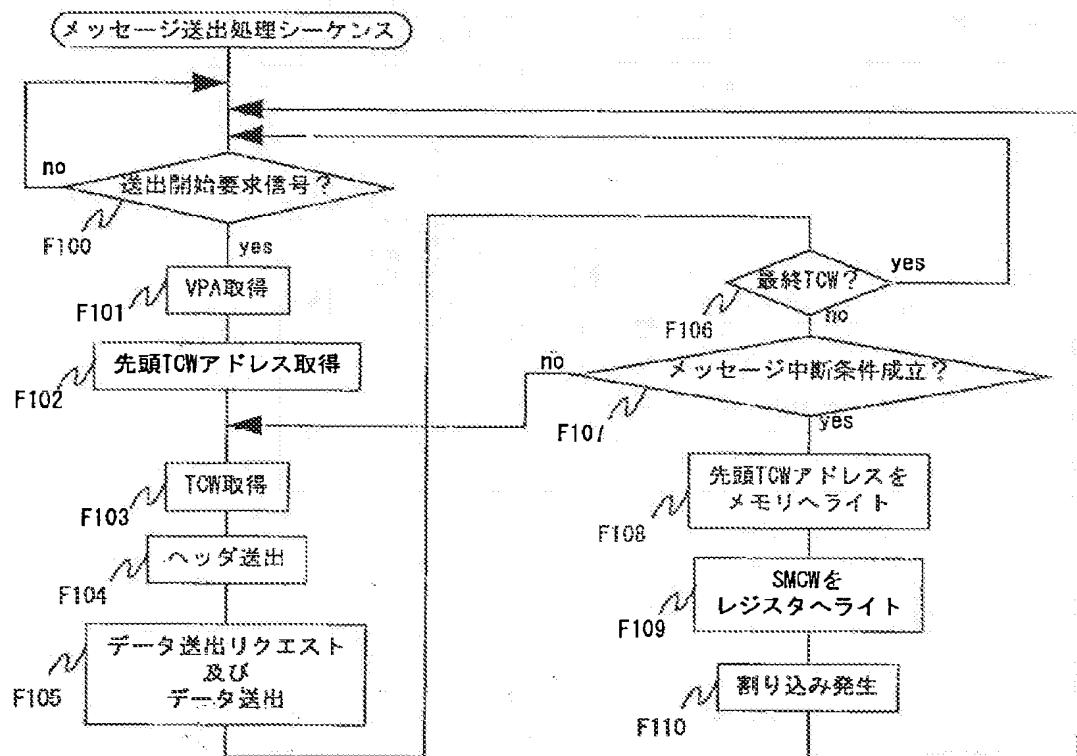
[图4]



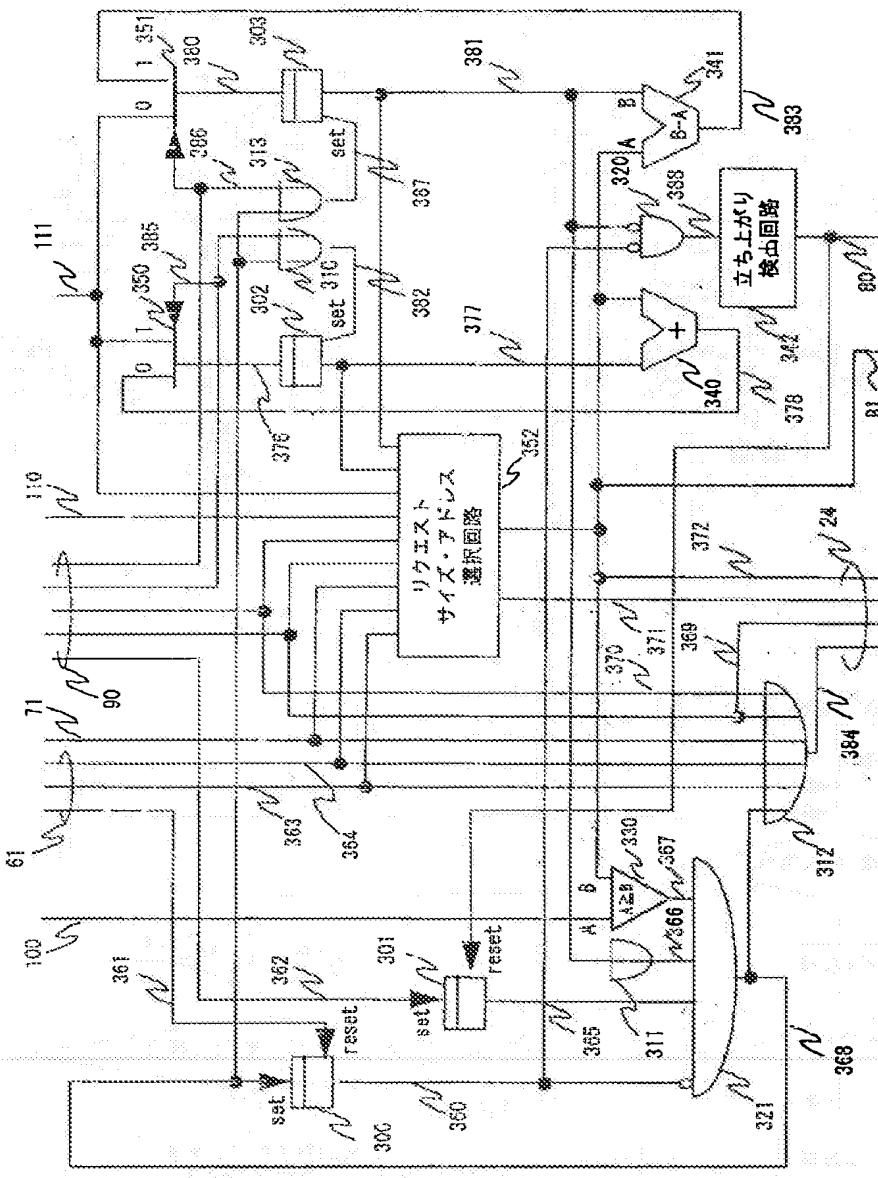
[図 7]



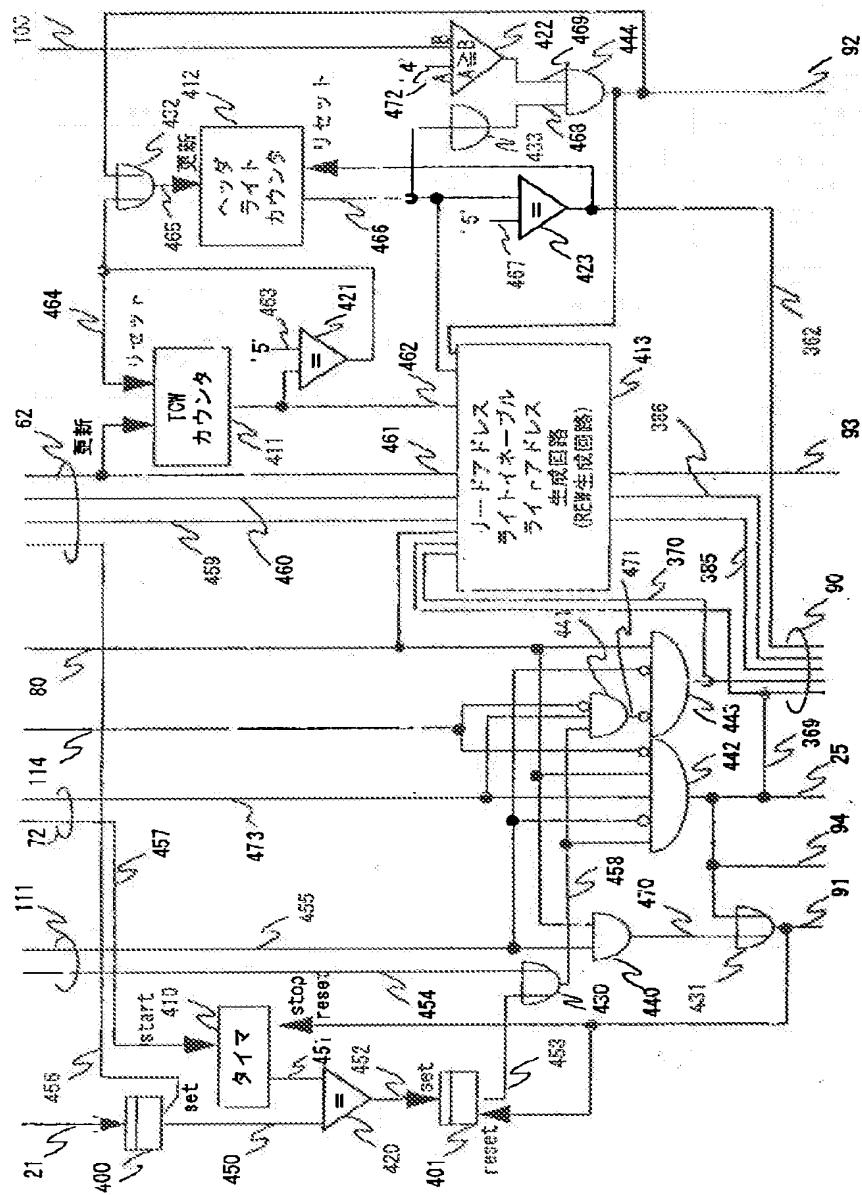
[図8]



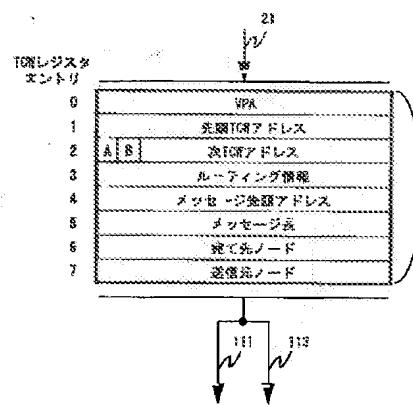
〔图5〕



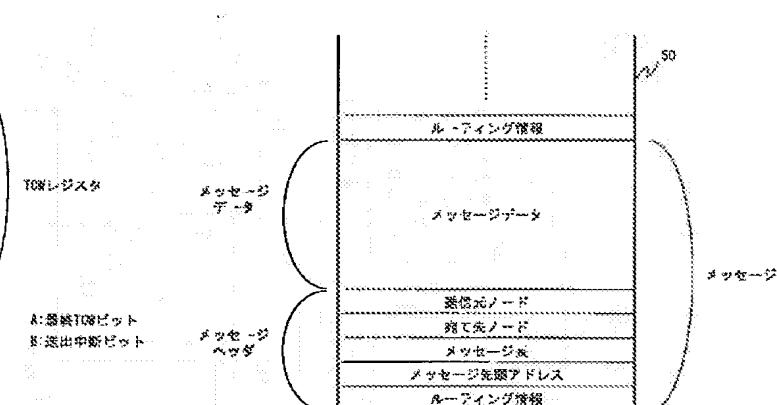
【図6】



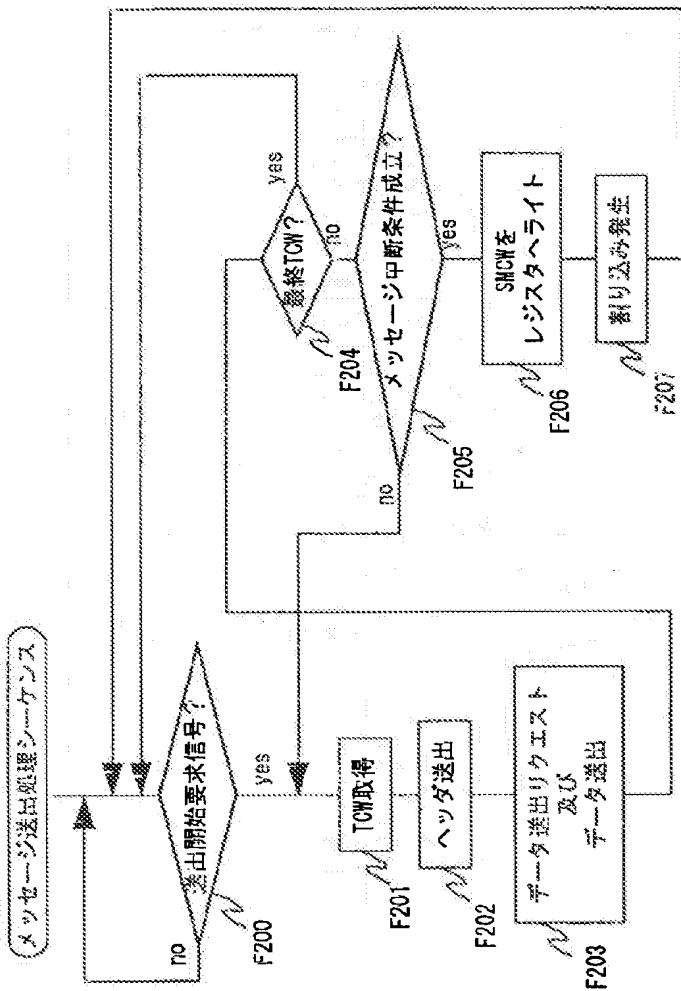
【図9】



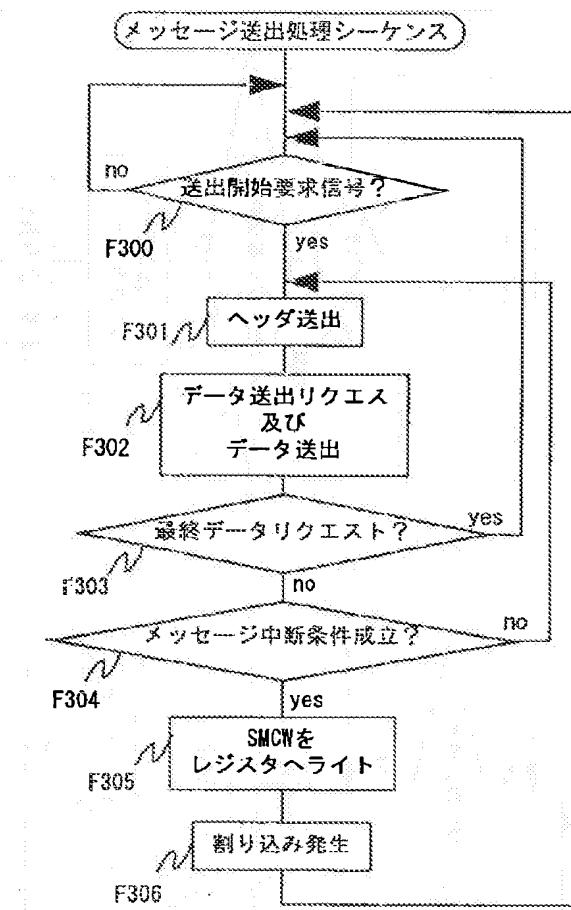
【図10】



【図11】



【図12】



【図13】

REW生成回路入出力対応表

INPUT	Output	ライトアドレス	ライト	リードアドレス	データ先端アドレス	データ後端アドレス
VPAリブライ信号	0	1	0	0	0	0
先端TCWアドレスシフト信号	1	1	1	0	0	0
TCWリブライ信号(TCWカウンタ=0)	2	1	*	0	0	0
TCWリブライ信号(TCWカウンタ=1)	3	1	*	0	0	0
TCWリブライ信号(TCWカウンタ=2)	4	1	*	0	0	0
TCWリブライ信号(TCWカウンタ=3)	5	1	*	0	0	0
TCWリブライ信号(TCWカウンタ=4)	6	1	*	0	0	0
TCWリブライ信号(TCWカウンタ=5)	7	1	*	0	0	0
ヘッダライト信号(ヘッダライトカウンタ=0)	*	0	3	0	0	0
ヘッダライト信号(ヘッダライトカウンタ=1)	*	0	4	1	0	0
ヘッダライト信号(ヘッダライトカウンタ=2)	*	0	5	0	1	0
ヘッダライト信号(ヘッダライトカウンタ=3)	*	0	6	0	0	0
ヘッダライト信号(ヘッダライトカウンタ=4)	*	0	7	0	0	0
整数TCWアドレスリスト	*	0	8	0	0	0
先端TCWアドレスリスト	*	0	9	0	0	0
リスタート信号	*	0	2	0	0	0

*: don't care

